Levels and empty categories in a principles and parameters approach to parsing

Kolb, H.P.; Thiersch, C.L.

Publication date:
1990

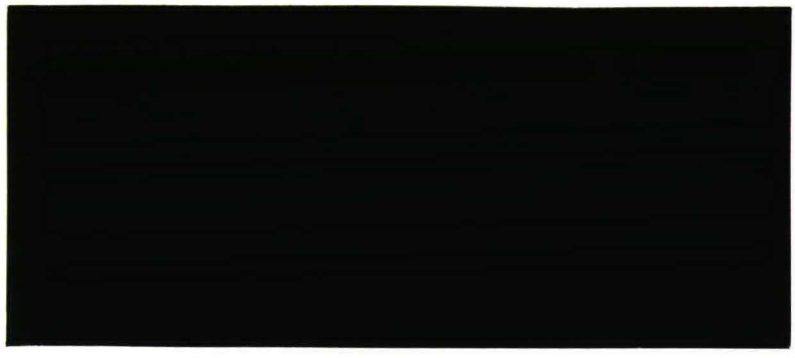
Document Version
Publisher's PDF, also known as Version of record

[redacted]

ITK

RESEARCH REPORT

INSTITUTE FOR LANGUAGE TECHNOLOGY AND ARTIFICIAL INTELLIGENCE
INSTITUUT VOOR TAAL- EN KENNISTECHNOLOGIE



ITK Research Report No. 19

July 17, 1990

Levels and Empty Categories in a Principles and Parameters Approach to Parsing

Hans-Peter Kolb
Craig Thiersch

ISSN 0924-7807

Levels and Empty Categories in a Principles and Parameters Approach to Parsing

Hans-Peter Kolb Craig Thiersch

Tilburg University

Abstract

This paper* discusses some basic problems of the implementation of a principles and parameter based linguistic theory. In the first part it outlines the distinguishing features of such a model, exemplified by its best known variety, Government and Binding (G/B) Theory (cf. Chomsky 1981, 1982, 1986a-b). Part 2 discusses some implications of preserving these properties for the design of a parser. Part 3 goes into several relevant issues in linguistic theory in more detail.

1 Introduction

G/B theory departs in several important respects from traditional models. Its main properties relevant for an implementation are

(1)
- The well-formedness of a structure depends on the interaction of general principles rather than specific rules.
- All structure is "projected" from the lexicon.
- There exists a "Universal Grammar". The grammar of a particular language is derived from UG by parameter setting.
- UG itself has a modular structure.

Let's look at each of these claims a bit more closely:

1.1 Principles vs. Rules

Most linguistic theories characterise the speaker's linguistic competence through a set of language specific rules, such as phrase structure rules or transformations, conflating the notions "grammatical construction" and "rule of grammar". These rules are highly specific and complex objects which generally describe the phenomena more or less adequately, but fail to capture a lot of important generalizations, both intra- and inter-lingual.

As an example, consider the following sentences and, at the risk of beating a dead horse, the (highly simplified) rules employed by early transformational grammar to analyze them:

*To appear in *Representational and Derivational Approaches to Generative Grammar*, Netter, K. ed. Dordrecht: Reidel.

(2) a. (John said that) Jill kissed Jack.

b. (John said that) Jack was kissed by Jill.

(3) • PS-rules:

$S \rightarrow NP\ AUX\ VP$

$VP \rightarrow V\ NP$

$(\ldots)$

• Transformation:

SD:	NP,	[+V,+Aux]*,		[+V,-Aux],		NP
	1	2		3		4
SC:	4	2	BE+EN	3	by	1

The passive transformation performs several operations simultaneously: It reverses the arguments of the verb, adjusts the inflection, adds two words, and in most versions, builds structure. But since the only connection between these operations is this not further analysable rule, it has nothing to say about the fact, that all the elements of "passive" also occur independently:

(4) • Movement of the logical direct object:

The destruction of Rome vs. *Rome's destruction*

• The passive morphology: *A beaten man...*

• The *by*-phrase for agent: *A book understandable by non-specialists...*

Moreover, there are phenomena with "passive" properties in many languages, but even in closely related languages they can't be described by the same rules, mostly for reasons that don't have anything to do with the construction in question, such as the different underlying constituent order in Dutch:

(5) a. (Kees zei dat) Jan Marie kuste.
K. said that J. M. kissed
'Kees said that Jan kissed Marie.'

b. (Kees zei dat) Marie door Jan gekust werd.
K said that M. by Jan kissed was
'Kees said that Marie was kissed by Jan'

Now assume that the notion of rule as sketched above is in fact a derived one, i.e. that grammatical constructions are not generated by specific rules, but are rather a function of the interaction of very simple, atomic declarative statements like the following:

(6) a. i. An argument has exactly one θ-role.
ii. An overt NP has Case.
iii. A NP can move.

b. i. A verb assigns a θ-role to each of its arguments.
ii. A transitive verb assigns Case to its object.
iii. Finite inflection assigns Case to the subject position.
iv. Passive morphology absorbs both the Subject- ("external") θ-role and Case-assignment to the object (Burzio's Generalization).
v. A preposition assigns Case and θ-role to its complement.

These statements (strongly simplified as they are) are enough to account for the relevant ("passive-like") properties not only of (2), but of (4) and (5), too: In the active cases of (2), (5) there is no problem. Both the subject and the object get a θ-role from the verb, and Case is assigned to the object by the verb and to the subject by the finite inflection (cf. **Jack (to) kiss Jill*, **Jan Marie kussen*). In the corresponding passive sentences no Case is assigned to the object and no θ-role to the subject. The only way to save the structure w.r.t. (6) is for the object to move to a position where Case, but no θ-role is assigned, i.e. the subject position, while the external argument can either not be realized at all (*Jack was kissed*, *Marie werd gekust*) or be provided with the necessary Case and θ-role by a preposition (whose special meaning links its complement to the absorbed agentive argument of the head, i.e. in this case the verb).

Obviously there is a fundamental difference between (3) and (6) only if the conditions of (6) apply blindly to any structure, i.e. if they are overall structural conditions on well-formedness. But if they do, then they are extremely powerful. Just adding the assumption that verbs like *seem* don't have an external argument, for example, is then sufficient to explain the following contrast:

(7) a. It seems that Jack is happy.
b. Jack seems to be happy.
c. * It seems Jack to be happy.
d. * Jack seems that it is happy.

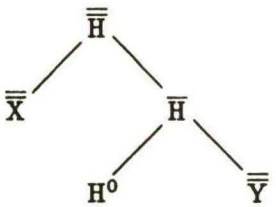
Moreover, if pretheoretic notions such as subject, object, assignment of θ-roles and Case, etc., are defined in general structural terms (*dominance*, *precedence*, *c-command*, *government*, etc.), it should be possible to do away with structure-building rules entirely, syntactic structure being completely determined by the interaction of the general conditions.

1.2 Projection from the lexicon

Only the conditions in (6a) state truly syntactic principles, while the statements in (6b) have a distinctly lexical flavour. In fact, the properties of lexical items seem to determine to a large extent the syntactic structure of a given string of words. This suggests a very strong stance of the lexicon in a principles based theory.

This extensive influence is expressed in the $\overline{X}$-hypothesis of GB-theory,[1] i.e. the axiom that all syntactic structure is endocentric and ultimately projected from a lexical head.[2] Hence a typical constituent structure is

(8)



(left-right order irrelevant
binary-branching accidental)

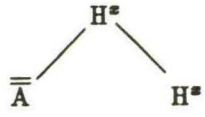
[1] This hypothesis is not specific to GB-theory. It dates back to early "EST"-times of Transformational Grammar (Chomsky 1970) and plays a major role in many current linguistic theories such as GPSG. In most formulations, however, it is used as a constraint on possible *rules* rather than as a direct structural condition.

[2] But cf. discussion in §3.

In a particular version of $\overline{X}$-theory we might now call the constituents in position $\overline{\overline{X}}$ Specifiers, and $\overline{\overline{Y}}$ Complements, whose actual realization would depend upon the lexical properties of the head.

The only configuration which is not lexically determined (but highly constrained by syntactic conditions) is an adjunction structure like

(9)



which, however, still obeys endocentricity.[3]

Note, however, that in a principles based approach the $\overline{X}$-hypothesis itself is to a large extent a derived notion. At least the argument structure of a constituent (cf. 8) will be determined by the structural conditions on Case and θ-role assignment.

1.3 Parametrizability

How, then, do we account for the differences between languages? Obviously it would be neither desirable nor necessary to construct a complete set of structural conditions for every language. The conditions in (6), for instance, seem to be sufficient to explain the characteristic aspects of "passive" phenomena (*inter alia*) in a large set of different languages. What distinguishes (2) from (5) is the fact that the internal arguments ("objects") of a verb appear on its right in English, but on its left in Dutch. As it is θ-theory that determines where arguments can appear in a structure, one way to express this is to say that verbs may assign their internal θ-roles to the right or to the left.[4]

Since θ-role assignment etc. is structurally constrained to government configurations, we can reduce this statement even further: Verbs may govern to the left or to the right, but they do it uniformly in a language. What we end up with is an example of a parameter which determines under what circumstances a universal condition on human language applies: Let's assume that Universal Grammar (i.e. the set of universally valid structural conditions) contains the statement

(10) A verb governs to the *dir*
where *dir* is a variable ranging over {left, right}

One of the differences between English and Dutch, then, is the choice of value for *dir*, the parameter setting.

A different type of parameter is exemplified by (11):

[3] Such an account, of course, renders the venerable sentential structure [S NP VP] malformed, but the idea that the sentence is actually a projection of the verb or of some inflectional element is neither new nor a special GB invention. Keeping these qualifications in mind, we will continue using *S* to label the sentential node.

[4] The matter is somewhat more complicated; cf. for example discussion of Chinese in Travis (1984), where it is suggested that θ-assignment is to the left but Case-assignment to the right, resulting in the variable position of objects to the left and right of the verb, with and without accompanying prepositions, respectively.

(11) a. (Jan zei dat)
Kees Marie de kinderen wilde helpen leren zwemmen.
K. M. the children wanted help teach swim
'Jan said that Kees wanted to help Marie teach the children to swim.

b. (Hans sagte daß)
Karl Marie die Kinder schwimmen lehren helfen wollte.
... swim teach help wanted

It is not hard to show that both sentences have identical underlying structures (cf. Evers 1975, Haegeman&vanRiemsdijk 1986) and that the main difference between Dutch and German w.r.t. this construction is the presence or absence, respectively, of an inversion mechanism operating on the verb cluster. Parameter setting in such a case would indicate whether a specific option is employed by a particular language at all.[5]

This feature of a principles and parameters based theory is, of course, not only theoretically (learning theory), but also commercially interesting: A principles-and-parameters-parser would by definition be a universial parser, i.e. it could be used for different languages by just exchanging the lexicon and fixing the parameters. Unfortunately the theory of parameters is, maybe not so surprisingly, still very much a *terra incognita*. There have been proposed interesting parameters accounting for various differences between languages, but the set is clearly far from complete, and even the discussion as to what counts as a parameter (as opposed to plain stipulation) has barely started yet. So, whatever the fate of the various GB/PP-parsing projects, the truly universal parser will remain in a visionary state for quite some time.

1.4 Modularity

GB-theory has a modular organisation in several respects:

First there is a methodological point: Lexicon, morphology, and syntax are claimed to form a self-sufficient unit which can be studied and explained independently of other modules of cognition. It is not obvious how this sense of modularity could have much impact on the design of a parser — unless modelling of human processing is part of the motivation of doing it, a topic to which we return.

Two other aspects of modularity, however, do inevitably influence the setup of an implementation:

The conditions of UG are not an unstructured set, but are grouped together in subtheories, modules. The canonical modules of GB are

(12)
- θ-theory, which, as we have seen, insures that arguments have a unique thematic role, and thereby determines where in a structure constituents with argument status may appear;
- Case-theory, indicating which elements assign and receive Case, thereby restricting the visibility of certain constituents;
- Binding-theory, dealing with constraints on referential dependency among constituents;
- Control-theory, which is concerned with the choice of antecedent for "understood subjects", i.e. with contrasts such as *Jack promised Jill [e] to leave* vs. *Jack persuaded Jill [e] to leave*;

[5] In this particular case, however, it may well be the case that we just have yet not reached a sufficient level of abstraction, i.e. the inversion facts may follow from the interaction of even more abstract conditions which in turn could be parametrised.

- Bounding-theory, which specifies certain locality conditions, mainly on movement;
- the Empty Category Principle (ECP), which imposes additional constraints on empty categories, especially on traces; and
- $\overline{X}$-theory.

These subtheories may share basic notions such as c-command which, in fact, plays a role in all modules, or government, relevant in θ-, Case- and Binding-theory, or co-indexing, pertaining to Binding- and Control-Theory, but they are conceptually independent in the sense that no module requires (the output of) another one to apply. They interact, but they don't depend on each other. Eliminating one or more of the modules doesn't cause the whole system to crash, it just makes it less restrictive. (Try that with a phrase structure rule...!)

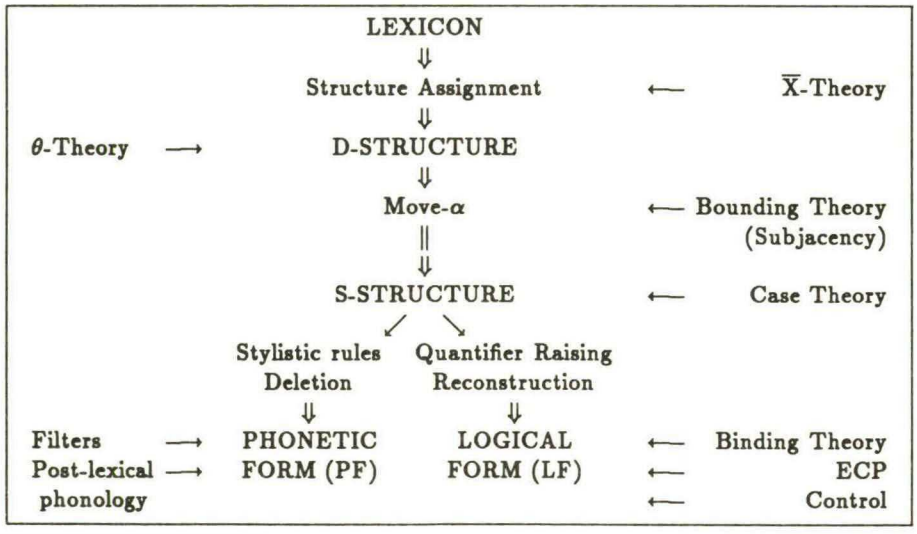
This is definitely a property of the theory one would like to preserve in a parser: It means that every module can be developed and tested independently and even changes in the final program can be confined to the module(s) they relate to.

The third form of modularity, the assumption of different levels of syntactic representation, however, spells trouble, as we will see shortly.

According to this hypothesis there exists a level of representation, D-structure, which contains the interface between syntax and the lexicon. It is a pure reflection of $\overline{X}$- and θ-theory as well as the lexical properties of the words employed. It is mapped via a general movement rule (Move-α, "move-anything-anywhere") onto S-structure, whose well-formedness is determined by the ECP, Case-, Bounding- and again θ-theory. This structure is in turn mapped on the one hand via stylistic rules, post-lexical phonology, etc. onto Phonetic Form, roughly corresponding to Surface Structure in vintage Transformational Grammar, and on the other via another instance of Move-α (Quantifier Raising) and other processes onto Logical Form, the interface to other modules of cognition, such as semantics, pragmatics, etc. Hence LF is also checked by a set of subtheories.[6]

We end up with the well-known T-model of the organization of grammar:

(13)



[6]The attentive reader may note some discrepencies between the text and the diagram; precisely which conditions apply to which levels is a matter of some controversy, and we take, for the most part, no stand here.

2 Principles based parsing

In general, a parser is a device which takes a string of words as input and returns some — preferably finite — time later either a structural description of that string or the message that it is not a well-formed string of the language. Highly simplified, it consists of three components — a grammar: G, a procedural interpretation of G: P, and a lexicon: D.

In a traditional parser, G consists of a uniform set[7] of rules — usually phrase structure rules "augmented" in some way or other — which constitutes a more or less adequate generative description of (some fragment of) some language L, D is an arbitrary list of <word, category> pairs,[8] and P can be understood as a procedural metagrammar[9] for G which uses G and D on a particular string of symbols S to decide whether S is a well-formed string of L, and, if so, assigns a structural description it.

While P determines the structural properties of G, its relation to L is very indirect[10] as is the one between D and L;[11] all relevant structural information about L is contained in G — which has to be written from scratch for every new language to be dealt with by the system. Insights into the universal properties of natural languages are not represented, not even representable, in the parser.[12]

In an abstract model of a universal parser based on a principles and parameters approach on the other hand the relations among G, P, D and L are very close, but considerably more intricate than in the traditional approaches: there is no set of rules which would define the structural properties of L. Instead, G consists of a set of universal conditions on possible structures of human language, i.e. UG. It determines the structure of a particular L in the same indirect sense in which the laws of statics plus the set of all communal building regulations in the world determine the structure of a particular building. P is a set of instructions (some of them parametrizable, in which cases P also contains the default settings) as to whether (under what circumstances) and how to apply the conditions of G.

P and G together define the notion "possible human language":

(14) Any language which — given an appropriate lexicon — is describable by a subset of G in a configuration permitted by P is a PHL.

Note, by the way, that the question as to what is the structurally unmarked language determined purely by P and UG is not well defined (although some of the work in the field seems to suggest American English as a good candidate...). Obviously it is the learnability of deviant settings for every single parameter, rather than the consistency of the whole set, which has to determine the initial parameter settings. Thus consistency, i.e. the existence of some "unmarked language", would be a rather uninteresting possibility following from nothing but the laws of pure chance.

---

[7] or a small number of such sets, c.f. the set of PS-base rules plus the set of transformations in classical transformational grammar.

[8] ...and where they aren't just this, as in GPSG, the additional lexical information is just a shorthand for more G-rules, as in all actual implementations of GPSG we are aware of.

[9] We disregard the differences between rule interpreters and compiled parsers as no theoretical issues seem at stake there.

[10] Generally P may restrict the type of L by imposing more or less rigid constraints on the weak generative capacity of G (cf. Tomita 1986), but most existing formalisms (ATNs, DCGs, Unification Grammar,...) have Turing power, anyway. Direct constraints on the structure of L can, however, be explicitly built into the formalism, as in Marcus (1978).

[11] Exemplified by lexical insertion (into full-fledged syntactic trees!) in early transformational grammar.

[12] This is one of the reasons why there are quite a few prototypes of parser schemes with rather impressive performance, i.e. Ps designed to work for any G written in a certain formalism (cf. Karttunen (1986), v/d Steen (1987) and many others), but so few parsers for natural languages which really deserve the name.

The individual differences between languages are encoded in a highly complex D which not only contains categorial information for the terminals of L, but also any idiosyncratic parameter settings. Thus the structural properties of some particular language L are a function of the interaction of P, G, and D. Given that P and G stay constant for all languages, the parser may be regarded as an interpreter of the lexicon rather than an interpreter of (language-)specific phrase-structure (and movement) rules.[13]

While anyone designing a parser for a rule-based model can draw on a large variety of well-understood methods,[14] it is obvious that retaining the relevant features of a principles and parameters based approach in a parser implies a major departure from the common scope as well as from the traditional techniques of natural language parsing. What, then, could a parser designed along the lines of such a theory look like?

As the answer to this question will largely depend on the final goal of an implementation, a short digression on possible motivations for such an enterprise seems in order.

2.1 Motivations and non-Motivations

If we ignore commercial applications for the moment, which are by definition uninteresting from the linguistic (though not necessarily from the computational) point of view because they depend on efficiency in the strong sense, which in turn implies heuristics, cutting corners, the supremacy of execution speed over theoretical aptness, we are left with two possible motivations: trying to simulate the "human parser"[15] or just writing a "theory development/testing tool".

On the background of the strong claims of G/B-type theories about the cognitive/psychological reality of Universal Grammar, modeling human processing seems to be an especially captivating option. Note, however, that, just as in early generative grammar, we are dealing with a theory of linguistic competence here, a fact that is particularly emphasized by its declarative setup. The common assumption (e.g. in Chomsky 1981) is that a child starts out with an innate "zero state" of linguistic knowledge, i.e. a representation of Universal Grammar with the default parameter settings, and some, again innate, language acquisition device (LAD). The LAD then uses *inter alia* linguistic experience to map the zero state onto a "steady state" of knowing and being able to use a particular language. While it is not obvious that the zero state is (part of) a parsing device at all, and the LAD still has very much the status of a "black box", the steady state has to contain, maybe even consist of, an efficient, robust, and highly specific parser. Is it this device that would have to be modelled in a psychologically relevant implementation? Apart from the fact that modeling the latter would set an end to any ambitions of universality, it also presupposes a matching computational theory of performance[16] accounting not only for the grammatical factors, but

[13] This property is claimed to be shared by quite a number of theories of grammar (GPSG, LFG, etc.). A discussion of these claims would exceed the scope of this paper. There is, however, one more major principles based approach, which shares it by definition: Generalized Categorial Grammar. (cf. Steedman 1987, Moortgat 1987, ...) A GCG with a general type raising rule, however, abandons (syntactic) structure: *all* relevant properties are lexically coded, structural considerations play no role in parsing (though they may play a role in the design of the lexicon). This avoids some of the problems discussed in the following paragraphs and allows for a strict left-to-right, "on-line" parsing strategy, but it is hard to imagine how strongly structurally constrained phenomena, such as the binding theory facts, could be integrated into such a system.

[14] cf., for example, Aho&Ullmann (1972) for a collection of algorithms.

[15] ... a motivation which, in fact, hardly anyone working on this matter fails to stress — no matter how implausible the actual system proposed may be in this respect.

[16] ... which, as Marcus (1978) has claimed, might even serve to explain some of the primitives of the theory of competence.

also for the rôle of semantics, pragmatics, and extragrammatical knowledge in general in human language processing.

But even if we neglect the complex and notoriously badly understood interaction between the various modules of cognition, two empirical questions have to be solved before anyone can embark on the task of simulating the human parser:[17]

(15) a. **What kind of a machine is the human brain?**
Is it a simple serial van Neumann computer, just incredibly fast and with no practical limits on space? Is it a parallel machine, maybe one whose processors are highly specialized special purpose inference machines? Or is the secret just "more of the same", including redundancies in the information structure, precompilation of patterns, heuristics, in short all the dirty things that a scientist carefully tries to avoid when setting up a theory?

b. **How are the modules/conditions/parameters reflected in the parser?**
Is the parser just an interpreter of UG and lexicon, as in our abstract model in the last section, i.e. is the learning of a language nothing more than the setting of parameters in the "database"? Or is it in fact the learning device which creates the parser by input driven deduction on UG and compilation of the theorems, in which case the parsing device proper would only very indirectly reflect UG? Or is it a combination of both?

These lists are certainly not exhaustive. But note that even if we had a conclusive description of performance phenomena, they could be mimicked in any computational model employing any random pair from (15a) and (15b). And this implies that even psycholinguistics is of only very limited help as long as there are no answers to these questions, which have barely been tackled up to now.

Hence, if we want to avoid pure speculation we have to fall back on the apparently least attractive option: the formalization and algorithmization of the theory. But as this is a necessary prerequisite of any further developments, the results stand a good chance to serve as starting points for more ambitious projects.

2.2 Investigating the model

Abstracting away, then, from the overload of "psychologically real" parsing, designing a Principles-and-Parameters parser reduces to the task of making explicit the trinity introduced in section 2, i.e.:

(16)
- design of the lexicon, including the problem of lexical derivation.
- axiomatization of G, i.e. explicitly formalizing the modules of UG. As there doesn't exist a complete, coherent and fully explicit version of a G/B-type theory, this is basically a linguistic problem, and a non-trivial one, too.[18]

[17] Marcus' Parsifal (1978) is a good example of an ingenious but premature step in this direction.

[18] Where the theory *is* explicit, however, formalization is not a very demanding task: Most of the notions are defined in a semi-formal way which can be easily translated into, for instance, a PROLOG definition. Consider as an example the following notion of 'c-command':

Node A c-commands node B,
if neither A nor B dominates the other and
if the minimal branching node dominating A also dominates B.

In order to write a working PROLOG definition we only have to make explicit the implicit reference structure 'A c-commands B in structure N':

```
c-commands(_A, _B, _N) :-
```

- design of P, i.e. the development and formalization of a parsing strategy, of a "driver". Given the theoretical objectives any principles and parameters based theory must meet, this can be done, though not finally implemented, independent from G.

It is this last point, where the basic problems of a G/B parser lie, even if we neglect the additional complications of on-line parsing and psychological reality. For the rest of this section we will therefore focus on the this topic.[19]

Remember that the G/B modules are considered as conditions on arbitrarily assigned structures, active on different levels of representation. In its pure form, then, the G/B parsing problem looks suspiciously hard, NP-hard[20] to be exact: Structural descriptions are easy to verify, but hard to get at. The undesirable consequences of this property become clear immediately when we ask ourselves what an appropriate driver could look like:

When starting a parse, all we have to work on is the input-string, roughly corresponding to PF in (13). According to (13) the only interface to the lexicon is located at D-structure. So if we want to stick strictly to the setup of the theory, we have exactly one option: We can generate well-formed strings starting at D-structure until we arrive at one that matches the input.[21] The disadvantages of such a procedere, however, are just too obvious. Without special provisions[22] we even end up with a halting problem, if the language to be parsed is not finite — and the infinity of natural language seems to be a pretty well established fact...

A much less obviously intractable strategy would allow access to the lexicon from PF, using the lexical information on sub-categorization and selection as well as the only module interpretable as an *overall* structural condition, $\overline{X}$-theory,[23] to assign structure to the input. That is, we would generate all $\overline{X}$ compatible structures, and use the other modules at the appropriate levels to filter the bad ones out. This leads to the following, "naive" parsing adaptation of (13):

```
        not(dominates(_A, _B)),
        not(dominates(_B, _A)),
        minimal_branching_node_dominating_A(_C, _A, _N),
        dominates(_C, _B).
dominates(_node, _node1) :- daughter_of(_node1, _node).
dominates(_node, _node1) :- daughter_of(_node2, _node),
        dominates(_node2, _node1).
minimal_branching_node_dominating_A(_C, _A, _N) :-
                node(_C, _N),
                branching(_C),
                dominates(_C, _A),
                not(node(_D, _N), branching(_D),
                        dominates(_D, _A), dominates(_C, _D)).
node(_node, _node).
node(_node, _tree) :- dominates(_tree, _node).
branching(_node) :- daughter_of(_daughter1, _node),
        daughter_of(_daughter2, _node),
        _daughter1 /= _daughter2.
```

The definition of *daughter_of* obviously depends on the preferred representation of syntactic trees, but is in any case trivial.

[19] For a discussion of a range of further linguistic problems, cf. §3.

[20] I.e., "as hard as" any problem in $\mathcal{NP}$, the set of problems solvable by a non-deterministic automaton in polynomial time. A formal proof of this property has been worked out in E.Ristad (1988).

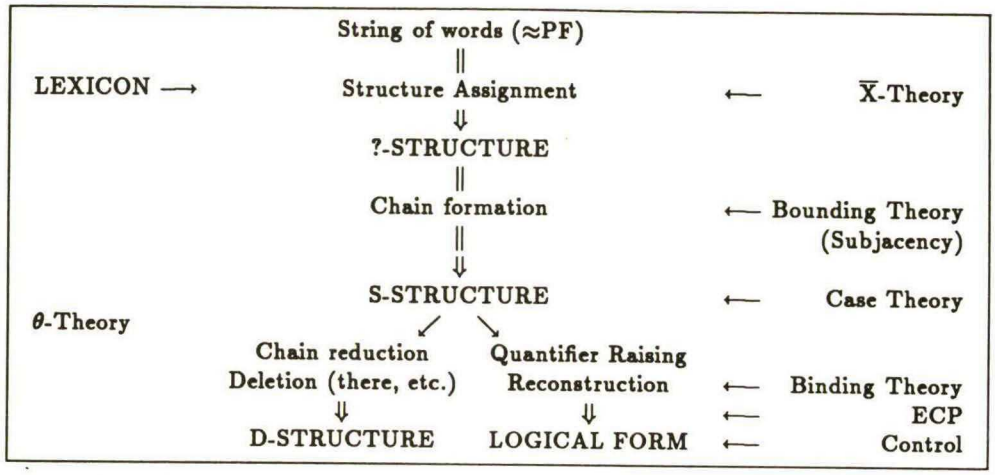
[21] This strategy is the starting point of Mark Johnson's (1987) G/B-parser.

[22] Restricting lexical access to the words occurring in the input string or making the width of the structure tree a fucntion of the input string would be obvious options to take — and pretty much the only ones, too.

[23] All other modules only involve relations between specific subparts of the structure, i.e. a binder and a bindee, a θ-assigner and an argument, a Case-assigner and an NP, etc...

Note that this step implies the theoretical, and linguistically relevant decision that $\overline{X}$-theory apply not only at D-structure.

(17)



This scheme introduces a new intermediate level of representation *?-STRUCTURE* consisting of a labelled bracketing possibly including empty categories, but stating no relations between parts except precedence and dominance. Move-α is split into two sub-processes, chain-formation deriving S- from ?-STRUCTURE, and chain-reduction deriving D- from S-structure.

But again we are faced with a halting problem: since ?-STRUCTURE may contain empty categories, without the modules constraining empty categories being available at this level, we end up with infinitely many possible ?-STRUCTUREs for any given string. Even if we employ *ad hoc* restrictions, such as allowing gaps only in argument positions[24] (which can be derived from lexical information), we get an intractable number of possibilties. A version of $\overline{X}$-theory along the lines briefly sketched above, for example, with Chomsky-adjunction permitted at any bar-level, but restricted to binary branching, will allow more than 35000 *different* ?-STRUCTUREs for the simple German subordinate clause *... dass der Karl den Hund schlug* ('...that Karl beat the dog').[25]

The conclusion is clear: Bad structures must be kept from being generated in the first place. But the generation of a faulty structure is only preventable, if all relevant conditions are checked as soon as possible already during structure building, i.e. they must be reinterpreted as conditions on structure assignment.

Aside from G/B-theory, the equivalence between the declarative and the procedural view on structural conditions is, for all practical purposes, a straightforward fact.[26] What complicates matters in our case, however, is that in the standard formulation of GB-theory the relevant conditions don't refer to the same levels, i.e. structures. Hence a procedural re-interpretation of the modules alone isn't enough. The interconnection of the levels, and their contribution to the grammatical well-formedness of a sentence, also have to be captured in an incremental way.

It is to date a matter of debate whether the relation between the linguistic levels is a truly derivational one, i.e. if there exist specific mappings to derive S- from D-structure and LF and PF from S-structure, with the modules purely functioning as filters on these derivations but too weak to define the levels by themselves; or whether the levels are, in fact, *constituted* by sets of modules, i.e. if they exist in parallel, connected only by the overlap of

[24] ...which would practically ban movement of adjuncts from grammar, provided this sort of movement leaves traces. See discussion in § 3.3

[25] Assuming, of course, that initially *no* conditions apply to Chomsky-adjunctions via Move-α.

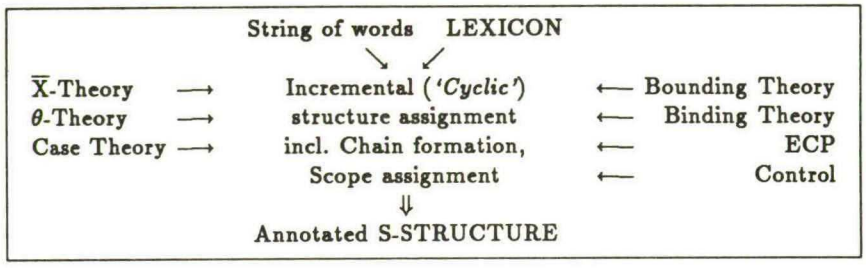
[26] cf. the interpretation of phrase structure rules as node admissibility conditions in GPSG.

the conditions that are operative on them.[27] It is obvious that the first view causes great problems for an incremental parsing model. In the absence of conclusive linguistic evidence for the necessity of a derivational approach, we will therefore adhere to the second, more declarative view.

We seem to be left with two options, then: On the one hand we could try to construct all levels in parallel. Such a setup would mean to parse what is in fact an S-strucure, but to use the interaction of modules and lexical information to predict for every substructure to be built a corresponding D-structure and LF, whose availability in turn determine the well-formedness of the S-structure.

Alternatively, we could compile the modules to constrain a single "annotated S-structure", which contains all the information of the original levels, leading to the following scheme:

(18)



This is a pseudo-alternative, however: Already in a canonical S-structure all properties of D-structure are trivially represented, except for the one fact that every argument has to start out from a θ-position. But this requirement[28] can easily be formulated as a condition on chain formation:

(19) For every chain C (which is a complete linear ordering $<$ on a set of syntactic nodes) and every node $n_i \epsilon C$, if n_i is a θ-position, then there does not exist a node $n_j \epsilon C$ such that $n_i < n_j$

i.e. if a chain contains a θ-position, then it must be its last element. It can be argued that this condition is just a descriptive statement while the assumption of a level of D-structure provides us with an explanation of the facts, but it seems reasonable to assume that the "backwards-Move-α" mechanism of the polystratal variant would use exactly the same constraint to construct D-structure.

An equally perspicuous relationship holds between any given LF and its corresponding S-structure.[29] What makes LF special in the canonical case is the fact that there may be more than one LF per S-structure.

Both the poly- and the monostratal option referred to above, however, treat a derivation (including LF/Scope assignment) homogenously, i.e. to different readings, whether purely structural or scope ambiguities, correspond to different derivations, and each derivation represents exactly one D-, one S-structure and one LF, either directly, or trivially derivable.

---

[27] In the standard formulations for example, specified by the Projection Principle.

[28] As observed by Goldsmith, ca. 1972, as quoted in Chomsky (1975), pp.115-17; cf. also Koster(1978), etc. Basically it boils down to traces (and phonetically realized morphemes like the preposition *by*) encoding the DS positions in SS.

[29] The intuitive difference in status between the two pairs reduces to the fact that we generally take S-structure for a true, in fact *the* primary structural level, while the status of LF is much less lucid in this respect. (cf. the presumable structural ambiguity of *Who did you promise to leave* vs. the scope ambiguity of *Some woman is loved by every man*; that the difference turns on structure vs. scope is less than straightforward in some cases; cf. discussion in §3.3, for example.) Cf. also footnote 38.

Hence the options are theoretically indistinguishable: If the levels can be constructed in parallel, then they can be conflated, too.

In any case we end up with one single condition on the combinability of two nodes:

(20) Nodes A^i and B^k can be combined to form node $\alpha = [_{A^j}\ A^i\ B^k]$
if B^k is licenced as a satellite of A^i in the configuration α,
i.e., if the structure α fulfills all modules relevant to B^k
(left-right order irrelevant; cf. discussion in § 3 as to the bar-level values of *i, j, k.*)

2.3 Locality

Such an approach reduces parsing to problem solving, where structure building is driven by a grammatical "expert system", a very desirable result given the declarative setup of the underlying theory. Its success, however, largely depends on the size of the local domains in which the modules can, in fact, be checked locally, in the linguistic as well as in the procedural sense.

Linguistically, a module M applies locally to some structure N, if N contains all the information needed to decide whether it is well-formed according to M.

It is easy to see that a module such as $\overline{X}$-theory can be checked locally in this sense within the smallest possible domain: any node. The same applies to θ-theory[30] under a bottom up strategy of tree construction.

(21) a. A structure N is well-formed according to $\overline{X}$-theory[31] iff
N is of category C and bar-level B and there exists a node D in N of category C and bar-level $B1$ such that D is a daughter of N and $B \geq B1$ or N is a terminal.

b. N is well-formed according to θ-theory iff
for every argument (= NP or complement clause) daughter A of N there exists a chain[32] (possibly of length 1) which contains A and a θ-position.

But not all modules have this property. Consider the following data:

(22) a. ...Jack to have seen Jill

b. ...Jack had seen Jill

c. ...Jack to have been seen *e* (by Jill)

d. ...Jack had been seen *e* (by Jill)

e. ...Jack to seem that Jill had seen him

f. ...Jack seems that Jill had seen him

[30] At least if we limit ourselves for the time being to the canonical formulations which have nothing to say about sentences like

Schlagen wollte Hans den Hund eigentlich nicht
beat wanted Hans the(acc.) dog actually not
'H. did not really want to beat the dog'

where the θ-(and Case-)assigner of the object has been preposed. In an A-chain $(x_1, x_2, \ldots, x_n)$, the distance between the Case-position x_1 and the θ-position x_n can presumably be indefinitely long; in §3.2 we look more closely at the locality requirement between links x_i and x_{i+1}.

[31] This is the most liberal version of $\overline{X}$-theory possible. More restrictive formulations can be obtained by adding the appropriate conditions to the consequens of the definition.

[32] ... of the proper specifications, the spelling out of which would lead us too far astray here.

When the S-node is being built in (22), the θ-module will accept the structures (a-d): (a, b) directly, and (c, d) via a non-trivial chain headed by *Jack* and including a θ-position. It will reject (e, f). A local Case-module designed along the lines of (21b) on the other hand would accept (b, d, f), but would not only correctly reject (e) but also — prematurely, at least — (a) and (c), which can be saved by becoming the complement of an exceptional case-marking verb like *believe*. The best we can do, in terms of locality, is to accept all the clauses, but to place a constraint on further processing, in this case the requirement for the resulting structure to get Case and transmit it to the subject-NP.

Note that while this example, as well as the entire discussion in this section, presupposes a bottom up parsing strategy, the locality problem is not restricted to, or even an artefact of, such an approach. As Chain formation is not an immediate option with a top down method, both Case and θ-theory cause problems when parsing *Jack*: If the clause is not embedded under an exceptional case-marking verb, Case theory has to restrict INFL to [+ tense]. In the case of θ-theory we even end up with a disjunctive constraint on the remainder of the sentence: either INFL assigns an external θ-role or the VP contains a gap in theta-position (i.e. some sort of GPSG-like slash mechanism).

A similar problem returns in the case of Binding theory:

(23) A structure N is well-formed according to Binding theory iff N has an indexing I such that

a. for every anaphor A in N,
if there exists a governing category G for A in N then A is bound[33] in G;

b. for every pronominal P in N,
either there exist a governing category G for P in N and A is free in G,
or A is free;

c. for every "referential expression" R in N, R is free.

The conditions in (23) can only give a definitive answer locally if the structure checked contains a governing category for every anaphor in it, i.e. the locality domain of (23a) is not any node, but rather any governing category, a much weaker notion.

But (23) also scores extremely low on locality in the *procedural* sense: Verifying it means exhaustive search of the complete (possibly deeply embedded) tree structure, one of the computationally most ineffective operations possible.[34]

Further complications arise with empty categories as in

(24)	a. Who do you believe [*e* to be funny]	*referential expression*
	b. Jack is believed [*e* to be funny]	*anaphor*
	c. Jack tries [*e* to be funny]	*PRO ("pronominal anaphor")*

not even the status of the NP *e* with respect to Binding Theory can be determined locally.[35]

Having noted some of the computational problems involved, let us now turn to the question of whether some revisions of linguistic theory might partly ameliorate the problem.

[33] X is bound if it is co-indexed with a c-commanding element (in argument position). X is free if it is not bound.

[34] While there are fairly obvious short-cuts for (23a, b), no such option exists for (23c).

[35] Cf. Kolb (1989) for additional discussion of a possible solution involving constraint propagation.

3 Some theoretical issues

As argued in the preceding section, in implementing a theory of the sort envisaged by the most common version of the principles and parameters approach, G/B theory, one immediately runs into a problem weeding out the multitude of possible structures[36] early on in the parse, caused in part by the structure of a theory which claims that certain conditions which license structure can only be "checked" by transforming, or rearranging, the structure. While we cannot review all of the arguments for and against a multi-leveled theory here, we would like to suggest some theoretical revisions which make a monostratal approach seem more promising, in that they limit still further the locality of the domain in which conditions must be checked. First, however, some preliminaries: beyond the general claim that language specific constructions result from the interaction of universal principles with language specific parameters (defaults) and lexical idiosyncracies, various versions of G/B theory (or the Principles and Parameters approach in general) make certain specific claims about the nature of Universal Grammar[37], the truth of which is an empirical matter, but which bear on the form of a "universal parser". In particular, the precise manner in which constituents are licensed will in some cases determine the locality of the licensing relationship as well as bearing on the question of whether a separate "level" of representation is necessary for the licensing to be checked.

Hence we begin by discussing the uniform constituent hypothesis and its relation to licensing in the next section (cf. §1.1 above); followed by a look at some areas where the proliferation of ambiguous structures can (possibly) be stemmed locally in building up S-structure: A-chains and Adjuncts.[38]

3.1 Uniform Constituent Hypothesis

One part of the Universal Grammar Hypothesis that has received renewed attention recently is the uniform constituent hypothesis (partially embodied in X-bar theory). As noted above in §1.2, all constituents project from a lexical head[39] according to universal principles and any idiosyncracies of constituent structure result from lexical properties plus (presumably simple)

[36] quite aside from the problem more often discussed in the parsing literature of deciding which of two alternate structures for a genuinely ambiguous string is the correct one on the basis of, say, semantic considerations.

[37] Several of which have been touched upon above in section 1.

[38] There are many other sources of proliferation of ambiguous structures, such as a Chomsky-adjunction (e.g. in scrambling languages and for extraposition), but to discuss them all here would take us beyond the scope of this article, and we limit ourselves to these two cases.

While it is in theory possible to imagine cases where an LF structure would influence the building up of an S-structure, e.g., *if* one were to assume Q-raising, one could imagine that for a particular string S-structure S_1 would permit quantifier movement to creating LF_1 giving the intended scope, whereas S-structure S_2 would create an ECP violation and hence the only reading would be a non-sensical, or unintended one; but this seems rather similar to the much-discussed attachment problem for PPs (cf. footnote 94 in §3.3.), which must rely on extra-grammatical factors to decide the upon the correct structure, and hence we have nothing to say about it here.

On the other hand, it *has* been suggested in recent work (Pollock 1988/89, Chomsky 1988) that checking conditions head movement in LF is crucial for determining the grammaticality. If so, this would be such a case. The latter article notes the computational problems this entails and hence takes a rather pessimistic view of the relationship grammar/parser. An alternative, at least for English, is sketched in Thiersch (1989; *forthcoming*).

[39] Of course, under a theory including "predication" as a basic concept, exocentric constituents could be allowed — e.g. the structure [$_S$ NP VP] would be licensed if VP is "predicated" of NP. We try in the following to justify the endocentric approach, for reasons which we hope will become clear in the following sections.

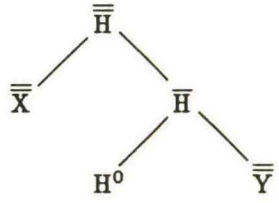
language-specific parameters. Of particular importance is the concept of "licensing"[40]: i.e., the idea that a constituent appears when and only when it is required by some independent module of the grammar. Hence PS-rules are otiose and, if we are lucky, nothing needs to be stipulated regarding phrase-structure as it follows from principles of independently motivated modules. In this spirit, let us take a particularly strong version of "X-bar" theory, and see how much can be derived:

(25)

- all lexical items[41] head a projection, including "little words" like conjunctions;
- that branching is strictly binary[42];
- that the bar-levels are represented by exactly two binary features, which indicate (1) whether a node dominates a Maximal projection of the head, and (2) whether a node is a Lexical item. (non-branching in the syntax)[43];
- all constituents have the same functional structure, consisting of a Specifier, Complement and (optional) Adjuncts, and these in fact have "typical" semantic interpretations, simplifying the semantic analysis of the constituents[44].

So for example, let us consider the canonical X-bar structure in (8), repeated here:

(8)

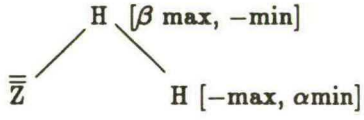


(left-right order irrelevant
binary-branching intentional)

where $\overline{\overline{H}}$ is [+max, −min], $\overline{H}$ is [−max, −min], and H^0 is [−max, +min]; $\overline{\overline{X}}$ is generally referred to as the Specifier and $\overline{\overline{Y}}$ as the complement. The representation of bar-levels as features can be exploited to eliminate vacuous (non-branching) projections, leading to structures like the following:

(26)

(a)



(b) $\overline{\overline{H}}$ (i.e., [+max, +min])

[40] Cf. Abney (1985). An early attempt to work out the details of the X-bar hypothesis understood as condition on possible structures appeared in Stowell's (1981) dissertation, and subsequent refinements and attempts to eliminate inconsistences have appeared in Muysken (1982), Pesetsky (1982), Gazdar and Pullum (1982), Thiersch (1985), Cann (1986) and Chomsky (1986b), Fukui and Spaes (1986) among many others.

[41] We leave open the (important) question as to whether strictly non-lexical items such as a hypothetical INFL have full projections. There is nothing crucial to parsing involved which is particular to INFL; cf. comments below in section 3.1.1.2 about bound morphemes and the instantiation of features.

[42] This leads to certain theoretical consequences which are discussed in the literature; cf. Kayne (1984).

[43] Cf. Muysken (1982) and Thiersch (1985) among others.

[44] Details in Thiersch (1985) and Cann (1986) among others.

Where the (a) structure is ambiguous, i.e. doesn't suffice to decide whether $\overline{\overline{Z}}$ is a Specifier or a Complement, and will be disambiguated by other modules such as θ-theory; (b) is, of course, a lexical word which functions syntactically as a maximal projection, such as *he* and *down* in *he fell down*. The structure in (26a) could be viewed as a filter on tree-admissibility, and (26b), as well as the full $\overline{\overline{X}}$ with Specifier and Complement are subcases thereof.

If adjuncts (modifiers of the head) appear at the $\overline{X}$ level, and only this level is recursive (the "traditional" view), this would indeed also be a subcase of (26a) and we would need to say no more about the phrase-structure. However, there are two reasons for relaxing this: Firstly, if we allow Chomsky adjunction to maximal categories, e.g., in "scrambling" structures:[45]

(27) a. Er meint, daß seinen Sohn ein toller Hund gebissen hat.
he means that his son a mad dog bitten has
'He says a mad dog bit his son.'

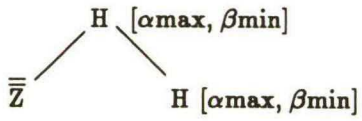
b. *presumable structure*:
daß [S [NP_{acc_i} seinen Sohn] [S ein toller Hund [e_i] ...]]

then at S-structure we have a violation of (26a). Secondly, we might want to allow adjunction directly to minimal categories, e.g., for the "unmarked" order of the arguments in German:

(28) ... daß er den Hund mit einem Stock geschlagen hat
that he the dog with a stick beaten has
'...that he beat the dog with a stick'

often referred to in the literature as the tendency for the direct object to be "left-peripheral".[46] We could relax the condition (26a) on phrase structure to allow identity operators, if properly licensed, yielding the structure in (9), repeated here:

(9)



where the $\overline{\overline{Z}}$ is *either* a "true" adjunct (modifier), licensed as indicated below in section 3.3, or a "scrambled" element, licensed as part of an A-bar chain.[47]

3.1.1 Licensing

Clearly not all constituents have the canonical structure, which after all is only a skeleton, or admissibility condition. From the above discussion, it should be clear why the constituent structure is different for different heads: the satellites appear only when properly licensed. Implicit in the preceding discussion, as well as in much of the literature,[48] is that the basic

[45] Nearly everyone does, although for different constructions – e.g. to VP in Chomsky (1986b), to S in Saito (1985), etc.

[46] It should be noted, however, that the issue is more complicated that this - cf. Czepluch (1988) for more complex data.

[47] Under a non-scrambling approach we would lose "parasitic gap" effects noted in Felix (1983), Bennis and Hoekstra (1985) and Thiersch (1985) which would arise under a scrambling analysis; cf. Huybregts and van Riemsdijk (1985) for an alternative. The issue may be a red herring under other analyses of case-marked θ-positions. Chomsky (1986b) contains a somewhat different proposal. Also see footnote 57 for a reinterpretation of the bar-level features...

[48] Cf. similar proposals of this nature in the work of Abney, Fukui, and Spaes cited in the bibliography.

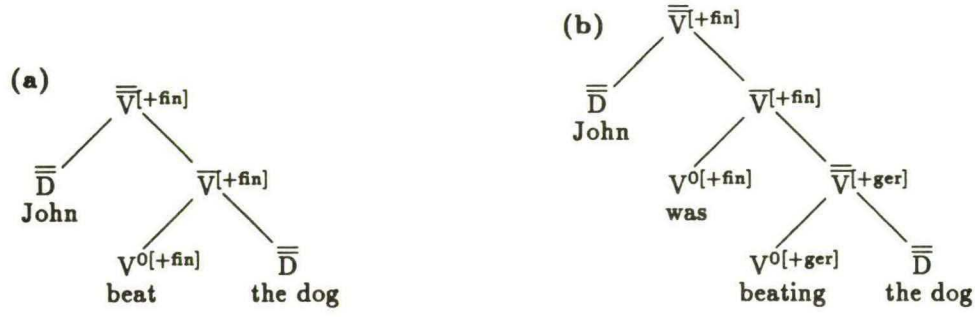
licensing relationship is between head and satellite.[49] Let us look at a specific proposal for the interpretation/licensing of the canonical Complement, Adjunct, and Specifier positions.

3.1.1.1 The Complement The standard assumption about argument positions like the complement is that assignment of θ-role licenses a potential position, and assignment of Case allows lexical material to appear in the position. (If no lexical material appears then the resulting empty category must be A-bar bound.)

This leaves other kinds of relationships between items which presumably stand in a *head:complement* relation unaccounted for. Hence let us assume that there are basically two kinds of heads,[50] namely true "content" predicates (such as main verbs, prepositions, nouns, etc.) and operators, such as modal verbs and complementizers: in the first case, the complement is an argument of the head in the usual sense, and in fact the one usually[51] referred to as the internal argument of the predicate; in the second, the complement is itself a predicate.[52]

3.1.1.2 The Specifier Let us further assume that the Specifier is always the "other" (i.e., external) argument, or, more precisely, receives the compound thematic role determined by the head and its complement (and adjuncts) as noted in the literature. In the case of a simple predicate head, it is the external argument, and in the operator case it is the undischarged argument of the complement.[53] For example, take the sentences *John beat the dog* and *John was beating the dog*[54]:

(29)



[49] Strongly implying, although not *a priori* necessitating, binary branching; see discussion of small-clause in § 3.2.2 and of double object constructions in footnote 85.

[50] For example, cf. Abney (1985); the idea itself is ancient, e.g. the notion of "full" and "empty" categories in Chinese grammar ...

[51] But cf. § 3.2 and especially § 3.2.3.

[52] Assuming here the analysis of auxiliaries argued for in Gazdar, et al. (1982) and elsewhere, where each is the complement of the preceding verb, and the verbal inflections are discharged in the same manner as Case. Cf. references therein. While this analysis might seem to create a proliferation of "maximal" projections which could cause problems for various parts of the binding/bounding theory, only certain of them will "count" (e.g. [+tense]).

[53] Cf. footnote 39 as well as Williams (1981) and others for details of a somewhat different view of the external argument. Alternatively, one can assume that the $(n\text{-}1)^{th}$ projection of the head is always predicated of the Specifier. The latter needs to be developed, but the idea seems promising; cf. recent work by J.-R. Vergnaud and several others.

[54] While we believe that the proper analysis for noun-phrases (traditionally NP) is really a DetP, i.e. the determiner is an operator which takes an $\overline{\overline{N}}$ as its complement (at least for some languages), nothing crucial to this paper hinges on the distinction, and we have hence left the alternating notations NP/DP or $\overline{N}/\overline{D}$ etc. unchanged. Cf. the articles by Abney, Fukui, and Spaes cited in the bibliography for discussion.

The verb *beat* assigns its direct object Case (e.g., accusative) to the right in both cases; but in (29a), the verb is [+tense] and hence can assign the subject Case, nominative, to the left, licensing this positon and allowing it to be assigned the thematic role of subject. In (29b) however, the verb *beating* is [−tense] and hence cannot assign nominative Case to its left. Thus the projection ends with the structure

(30) $[_{V[+max]}\ [_{V[+min]}$ beating$]\ [_{DetP}$ the dog$]]]$

since no specifier is licensed and therefore there is no "subject". In the next projection, however, *was* is [+tense] and can assign nominative Case, licensing the specifier position, but has no θ-role of its own to assign. Its lexical specification includes a complement of, say, [+V, +ING], which is fulfilled by (30) above. Since (30) is an open predicate, and the verb *be* links the nominative NP to the open argument position of its complement, the DetP *John* is assigned to the "dangling" argument position.[55] (The [+V, +fin] head of course has an additional semantic contribution, minimal in the case of BE, but considerable in the case of modals, for example). It is easy to see how this applies to the other auxiliaries, embedded auxiliaries (which are subcategorized by their complements just as "main" verbs are by *their* objects), and to other complement types: *John is very tall*, *John is in the kitchen*[56] as well as, possibly, in other kinds of projections: e.g. the DetP where the DET *'s* assigns Case: *John 's beating the dog*. This has some consequences for the formation of A-chains and will be considered below again in § 3.2.[57]

Note that in (29) there is no separate INFL projection; it seems unnecessary in light of the above discussion, and stems from a more general consideration: if we consider elements such as DET, INFL, COMP to be operators in the sense of Abney (1985), Fukui and Spaes (1986), then in some languages they will be realized on the surface as independent morphemes (the article in English, German, or INFL in Warlpiri) and in other cases

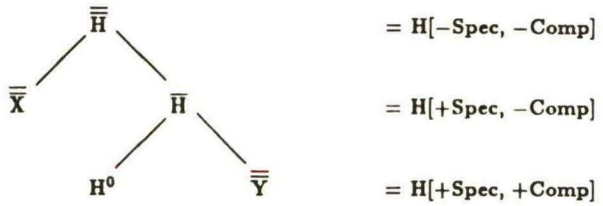
[55] Cf. details in § 3.2.

[56] Assuming the preposition in this example is a two-place predicate. This suggests that what characterizes adjunct modifiers such as

the man [with the hat]	P-max
a student [sick of the exams]	A-max
a man [to fix the sink]	V-max
the cat [which Fred saw]	C-max

is that they are all *syntactically* open predicates with the unsatisfied argument position in the Specifier position and is coindexed with an appropriate referent within a certain domain, an idea which will will be taken up and discussed in more detail in § 3.3

[57] Note that under this interpretation the features no longer have the interpretation usually associated with them in the works cited above (such as "projection", "word", "lexical", "maximal"), but come close to meaning something like [±Spec, ±Compl]:



i.e., [+Comp] would mean the head still needs to discharge a complement; [−Compl], that the H (or its projection) either *has* a complement (which may be a Case-marked empty category) or does not subcategorize for one, and the "condition" on X-bar structure is simply the familiar one that, if there are both an internal and external argument, the "internal" one must be discharged first; cf. discussion in Burzio (1981/86) and § 3.2. Since adjuncts (see § 3.3) are not arguments of the H-projection, they fall outside of this system and hence can be freely adjoined if they are (a) properly licensed and (b) violate no other constraints (including directionality of headedness). We will nevertheless continue to refer to the features as [max] and [min].

as a bound morpheme (e.g., the [definite] article in Skandanavian, Bulgarian, or INFL in many European languages.) Presumably its effects can be redefined as effects of the features such as [+tense]. That is, features which themselves play a crucial role may nevertheless be instantiated independently in some languages and as inflections in others.[58]

It should be noted that in addition to the constructions discussed here and in the references which do adhere to the canonical structure, there are many which remain problems: for example, clitic and clitic-doubling constructions, or languages with purported multiple WH-extraction, like Polish and Romanian.[59] Clearly we cannot discuss all such cases in a short article.[60]

Let us first look more closely at the status of empty-categories which result from "movement" in the usual sense and are cases of A-binding, where we will suggest that they are more locally bound (and licensed) than empty categories resulting from A-bar bound movement. We then turn our attention the proper treatment of adjunct modifiers, another source of structural proliferation.

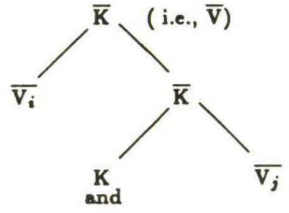
---

[58] In the case of *English*, as opposed to, say, Icelandic (Holmberg 1986) or French (Pollock 1989) — it is far from clear that there is head-movement into an INFL position — evidence consists primarily in "do-support", which in Subject-Aux inversion is movement into COMP, as in other Germanic languages; before *so* where it applies equally to other inflections (..., *and John might be doing so, too*); and before certain cases of *not*, the only real evidence. But the latter has a rather more plausible non-movement analysis. For discussion of the evidence, cf. Thiersch (1989; *forthcoming*).

[59] The latter is particularly instructive, as the framework described in the body of the paper does not countenance $[_{Spec(COMP)}\ WH_1\ WH_2\ \ldots WH_n]$; Toman (1982) and Cichocki (1983) have shown that in the Slavic case, only the first WH is in COMP and the rest are "scrambled" to the front; Comorovski (1986), on the other hand, shows this is not the case in Romanian, as all WH-words are subject to long-distance extraction. Her solution, however, namely adding an *ad hoc* PS-rule for Romanian COMP is clearly unacceptable in the Principles and Parameters approach in which language-specific PS-rules are eliminated in favor of licensing and parameters. In her paper there are a myriade of tantalizing details suggesting more principled avenues of approach...

[60] Finally, we should note for completeness that there are two cases deserving special attention: first, if we assume that the function of some operators is (among other things) to change the categorial status of their complement, e.g., the argument $COMP_a$ takes an assertion ($\overline{\overline{V}}$[+tense]) and allows it to function as an argument, or the relative $COMP_r$ which allows its complement to have an empty argument position so that the assertion can function as an unsaturated predicate (modifier). (That they are different is easier to see in languages where they are morphologically distinct, e.g. Bavarian *daß* vs. *wo*.) The Specifier of COMP in these cases, of course, does not get a compound θ-role as described above, but just the θ-role assigned to the A-bar chain.

The other case is coordination: assuming the analysis of coordination suggested elsewhere (for example in Thiersch 1985), the conjunction, an operator, is a *defective* head, underspecified for features, which unifies its categorial and bar-level features with those of its satellites (corresponding to the usual Specifier and Complement positions). Hence coordinated $\overline{\overline{V}}$'s form a $\overline{\overline{V}}$ and N's form an N, etc. So, for example, in



the whole projection is also a $\overline{V}$. Here the positions do not have the canonical interpretaton just alluded to, and furthermore can iterate. Aside from the details, discussed elsewhere, this is basically the minimal assumption we need to make about coordination.

3.2 A-chains

In § 3.1.1.1 above, we suggested that, in the case of a complement which is not a semantic argument of the head (i.e., grammatical object) but rather an open predicate (such as the complements of modal verbs or the copula) the undischarged argument of the Complement (generally the "external" one), or rather the corresponding compound θ-role, is assigned to the (potentially unrealized) specifier of the next head:

(31) $[_{\overline{\overline{V}}_{[+fin]}}$ John$_i$ is [PP θ_i [in the kitchen]]]

Hence, in a sequence of embedded structures (Complements) with unrealizable specifiers, the θ-role of the potential "Specifier" or external argument position of each Head, can be assumed to be the undischarged one from *its* respective Complement:

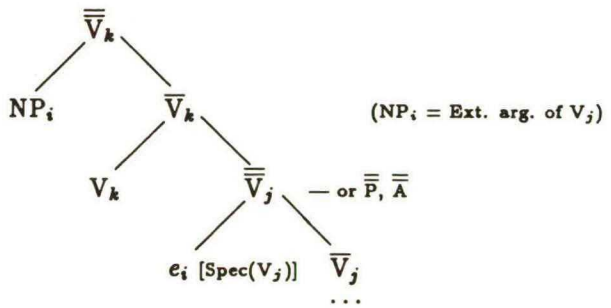
(32) [John$_i$ might [e_i have [e_i been [θ_i beating the dog]]]]

in this case the external argument of *beating*. (The e_i's only serve to indicate how the chain is formed and are not intended to indicate empty categories in the G/B sense.[61]) This looks suspiciously like the cases discussed in the literature under the rubric of "A-chains": a category "moves" from its base position (a θ-position) because there is no Case assigned, in a series of steps through θ-less and Case-less position until it ends in a Case position:

(33) John$_i$ seems [$_\alpha$ t_i to have been beaten t_i]

where the locality of the operation is guaranteed by the binding theory, which requires that the empty categories each be bound in their governing category, appropriately defined. Since one needs to say something about the former case (e.g., auxiliaries) anyway, one might ask whether they are both part of the same process, in which case A-binding is automatically strictly local, since it is always mitigated by a head. We would like to maintain a structural requirement limiting the linking of the "A-chain" arguments to the strictly local structure:

(34)



where V_k is the operator which transmits the θ-role.

This has both advantages and disadvantages: On the one hand it captures both the intuition and the facts about the strict locality of A-binding, and loses the parallel to the binding of overt anaphors, which seems spurious in view of cases like *They believed that picture of each other were on sale* which have no parallel in the case of empty anaphors. That is, the extensive discussion as to why there is no "super-raising" is spurious.[62]

On the other hand, there are (at least) two problematic considerations: One is the case of infinitival complements as in (33) above, or

[61] Although one might countenance a revision of the theory in which they were. In fact, there is some evidence that these Specifier positions do have empty category status and are part of a chain in some sense: cf. discussion in §3.2.1 below as well as recent work: e.g., Koopman and Sportiche (1988), Sportiche (1988).

[62] as are some ECP cases, as well. See ensuing discussion, and § 3.3.1 as well.

(35) John tried [PRO to leave]

where the complement has often been assumed to be an $\overline{S}$, i.e., $\overline{\overline{C}}$, preventing Case assignment to the position of the PRO/trace. If it is true in the case of *seem*, as well as other cases discussed below, that the arguments are really transmitted by intervening heads in the same way as in (31) and (32), then we need to assume that α in (33) consists of only one projection, too, and not a vacuous double, i.e., $\overline{\overline{INFL}}$ or $\overline{\overline{V}}$ inside $\overline{\overline{C}}$,[63] as we would not otherwise have strict locality for the passing of the θ-role (cf. details in §3.2.2). However, as has been noted in the literature, these putative $\overline{\overline{C}}$'s nevertheless have to be transparent for government in certain cases like (33) or else an ECP violation would result from the trace of movement not being governed, but not in the other case: (35). Furthermore, we have the case of *believe*, discussed below, where Case *is* assigned to the DP, and something like $\overline{S}$-deletion must be stipulated. Rather than this unsatisfactory state of affairs, let us observe that two things must in any case be specified idiosyncratically in the lexicon: the assignment of Case[64] and categorial subcategorization: e.g., whether a verb takes as complement a DP, CP or both. Crucial to the latter is that we differentiate the categories more finely on the basis of their feature composition; for example, the complement of the auxiliary verb *have* is not just any VP but the maximal projection of [+V, −fin, +part.]. Suppose we assume that *seem* no more assigns Case here than does the auxiliary verb *might*, so the $\overline{\overline{C}}$ is not necessary to block Case assignment.[65] That is, there are some (Case-assigning) verbs that take a CP type of *to*-projection,

(36) a. He knows that fact.
b. He knows [what$_i$ [PRO to do e_i]]
c. * He knows John to leave.

some which do not

(37) * It seemed [who to visit]

and some which take both

(38) a. John discovered [how [PRO to solve the problem]]
b. John discovered [the problem to be unsolvable]

Evidently *discovered* assigns Case (*John discovered the solution*), which is assigned in (38b) but blocked by the intervening C-projection in (38a); cf.

(39) * John discovered why the problem to be unsolvable

It is then no longer necessary to invoke the ECP to rule out the A-chain in (40b):

(40) a. [the problem]$_i$ was discovered [$_S$ [e]$_i$ to be unsolvable]
b. * [the problem]$_i$ was discovered [$_{S'}$ why [$_S$ [e]$_i$ to be unsolvable]]

[63] Unless we are willing to except a dummy phonetically null COMP which acts as a quasi copula, a move which seems unnecessary in view of the discussion in the text ...

[64] Obvious for so-called lexical case, e.g., Dative vs. Genetive; cf. further discusssion in §3.2.3

[65] If Burzio's hypothesis, $-\theta_{ext} \rightarrow$ −Acc., is true, *seem* could not possibly assign object Case. Modals in certain languages seem to present a problem: *Das*$_{[acc.]}$ *kann er doch!*.

as the chain is simply not local in the sense of (34).

The raising cases like (33) would then in fact work the same way as (31), but the "empty category" in the specifier position would be the same as whatever is indicated by the θ_i's and e_i in (31) and (32). That is, *seem* will also function rather like a modal, assuming that *all* modals take projections with propositional content and it is only chance (parametric variation) that some of them subcategorize for a $\overline{\overline{C}}$ as well as a $\overline{\overline{to}}$, or $\overline{\overline{[+V,-fin]}}$[66]: for example, in English *want* is infelicitous with *that S* complements, but acceptible with *for* infinitivals (with or without realised subject, depending on Case assignment); German *wollen* on the other hand *does* take a *daß*, i.e., *that* complement, but only a bare infinitival (*Er wollte gehen*; **Er wollte zu gehen*); nevertheless all of the complements are propositional in nature, involving a full argument structure.[67]

Before we attempt a partial formalization of these ideas, we need to discuss some of the problems related to the passive construction.

3.2.1 Passive

Suppose for the moment that A-chains are *always* constructed by linking the available argument of the Complement with the Specifier as suggested above, rather than by Move-α as is the case of $\overline{A}$-bar chains. As noted above, in the normal case it is the "external" argument which is linked, as in *John was beating the dog.* In the case of the past-participle, however, no Case is assigned to the complement position.[68] If its projection is embedded under an auxiliary which assigns accusative Case, such as *have*, the Case is exceptionally passed down, and the object can be realized *in situ*; if not, then it cannot be phonetically realized[69]:

(41) a. Fred *had* beaten the dog.

b. The dog *was* beaten — .

This can also be seen from the fact that the past participle always has the passive interpretation in isolation

[66] That is, the maximal projections of *to* and [+V,−fin.].

[67] This suggests, by the way, that what is distinctive about the *try* (control PRO) cases is that the *to*-projection has an empty element which may be *referentially* independent, and it is not θ-independence, as has often been assumed, which distinguishes PRO. If we consider cases like

i. John$_i$ wanted [e_i to sing a song]

ii. daß Hans$_i$ [e_i ein Lied singen] wollte (*lit.* '... a song sing wanted')

iii. daß Hans$_i$ [e_i das Lied singen] kann (*lit.* '... a song sing can')

iv. John$_i$ can [e_i play the piano] =

v. John$_i$ is able [e_i to play the piano]

it is hard to imagine any difference in the θ-roles in the PRO versus the modal cases; if there are two θ-roles for the coindexed categories in (i) and (v), then there surely are in (ii)–(iv). Referential independence, however, is suggested by the contrast

vi. Fred wanted PRO$_{arb}$ to meet at 5pm.

vii. * Fritz wollte e_{arb} um 5 Uhr zusammenkommen.
(*lit.* '... wanted at 5 o'clock together come')

[68] There seems to be no justification for assuming that there are two distinct forms, the past participle and the passive participle. Unlike, for example, the simple past tense and the past participle, where some forms are identical and some are not, and hence the language learner could posit different forms, the former are *always* identical. This has been noted elsewhere, e.g. Höhle (1978).

[69] Cf. discussion in Thiersch (1985), for example, and for a suggestion of how the Case is exceptionally assigned, Vergnaud (1985). Note that *accusative* Case is necessary to realize the internal argument in this position, and the Case (nominative in English, instrumental in Russian, etc.) assigned in predicative constructions doesn't count, for whatever reasons. Cf. Vergnaud (*op cit.*) for discussion.

(42) a. [NP the boy [beaten (by the police) yesterday]]

b. * [NP the boy [beaten the dog yesterday]]

cf. [NP the boy [beating the dog]]

Hence if we imagine the arguments to be on a stack, with the *internal* argument on top, i.e., with the highest priority,[70] then *it* is the only one available for being linked to the subject in this case, hence the passive construction. If Burzio's hypothesis is true,[71] then the *have* in example (41a) has assigned its accusative Case (exceptionally), and therefore must assign a θ-role to its subject position; presumably the possessive role is not available (different *have*) and so the "next" one available is the external one of the participle. Since the accusative Case is presumably being assigned by the *have* and not the participle, the object does not need to be governed by the participle and could just as well appear in the Specifier position of the participle, giving the "small clause" reading[72]:

(43) John had [[his car]$_i$ stolen e_i]

This gives us a possible answer to the question posed in footnote 61: are the e_i's in each intervening (auxiliary) verb in (32) real positions which form part of a chain? In (43) the presence of *his car* in this position evidently blocks the realization of the external θ-role of *stolen*, but the *have* still needs to discharge a θ-role, and the only θ-role left for the *have* to assign to *John* is the "patient/benefactive" one. In other words, the intermediate e_i's apparently must be unoccupied for the θ-role to be passed up. Thus we may be able to maintain the structural requirement for the linking of the "A-chain" arguments suggested in (34). In this way we can generalize the account of auxiliary verbs and A-chains above to cover passive as well, by referring to the "first available non-discharged argument".[73] However, there is a problem for all accounts: assuming that the external θ-role is unavailable outside of the local domain (bottom-most V-projection), the stack-model taken literally (as in the next section) would imply that the external θ-role is *truly* unavailable in passives. Yet it seems to be "picked up" by the by-phrase, and, contradictorily, available nevertheless. We present a solution in § 3.3.2 after examining the status of adjunct (modifier) phrases.

Before dealing with some potential problems, let us look more closely at some details of Case- and θ-assignment, and attempt to (partially) formalize this.

3.2.2 θ-linking and Case assignment.

Let us suppose for concreteness that we have associated with each head three *ordered* lists[74]: a Case-list, θ-list, and a category (subcategorization) list, each consisting of not more than

[70] Cf. footnote 57 where it was suggested that the features correspond rather directly to the argument structure.

[71] I.e. the contrapositive of $-\theta_s \rightarrow$ −Acc.: If a verb assigns Case to its object, it must assign a θ-role to its subject. Note that while this may play a role in the construction discussed in the text, passive is no longer a subcase, as it results from removal of Case, under this analysis, and the absense of subject θ-role then follows as discussed in the text. Perhaps the arrow goes both ways, although there is some evidence that the generalization may be spurious; cf. Napoli (1988) and the discussion in §3.2.3.

[72] Probably *should* be considered a small clause. Cf. subsequent discussion, esp. § 3.2.2.

[73] Assuming, as is standard, that an argument which gets Case is discharged and must be realized by a true syntactic position – either phonetically realized or A-bar bound. Cf. further discussion of details in §3.2.2 below.

Recall that the other case of argument reduction, true intransitives such as *eat* must be done in the lexicon, as they are highly idiosyncratic, as has often been noted in the literature: *John ate yesterday* versus **John devoured yesterday*.

[74] Lists in the computational sense, i.e. pushdown stacks, not just enumerations.

two elements.[75] So, for example, the θ-list for all forms of BEAT would be (θ_{int}, θ_{ext}),[76] whereas the θ-list for an auxiliary verb, such as HAVE would contain a variable in the second position: (θ_{int}, $\_\theta_{ext}$).[77] The Case-list for *beats* would be (Acc., Nom.), for *beating* (Acc., $\_x$) and for *beaten* ($\_x$, $\_y$).[78] Further, we need to indicate which elements are Case-requiring (DetP) and which are Case-resistant (CP). Suppose for the sake of argument, that part of the Case specification of a DetP contains a variable $< +, -, \ldots, \_C, \ldots >$ in the place where a Case-resistant item would contain a constant; consequently the Case resistant element will not unify with a specification containing Case and hence can only appear outside the domain of the Case-assigner, as required; the "Case-filter" can be simulated by requiring that the variable $\_C$ be unified with a constant (from the Case-assigner) at some point.

We can now translate the simplex cases discussed earlier into this formalism: in *John beats the dog*, for example, the partially complete Case feature list of *the dog* is unified with that of *beats* creating a list of constant features (no variables) and the (internal) θ-role can be popped from the list. In *John is [* e_i *beating the dog]*, the external θ-role is "next" on the list (the internal one having been already discharged by the accusative Case) and is available to be associated with the specifier position of *beating* and can be unified with the external argument of *is*. Notice then what happens to the Case in the so-called small clause constructions:[79] If we unify a DP, *Fred*, with a non-Case assigning element, like *clever* or *in her hands*, then the variable $\_C$ in the DP remains a variable and is unified with the Case specification in the resulting constituent, which then also becomes a Case-requiring entity.[80] If this projection is then further combined with a Case-assigning element then the variable $\_C$ can be unified with a constant, and the Case-filter is satisfied by unification, as the DP Fred now also contains the constant, as in

(44) a. Susan considered Fred clever.

b. Susan had Fred in her hands.

c. Susan believed Fred to have left.

The auxiliary verb *had*, however, is different in that its internal Case, for whatever reason,[81] may be unified not only with the specifier (external) Case, but the internal Case as well. That is, if we unify *beaten* with an (accustive) DP *John*, so that [$_{VP_{part}}$ *beaten John*] itself becomes a Case-requiring entity, as indicated above, the accusative Case of *have* may be unified with the internal Case specification, giving *had beaten Fred*. We can see this is exceptional by

(45) a. * Susan considers beaten Fred

b. * Susan believes beaten Fred

As noted above, we need to require in any case that the next available θ-role is added to the θ-list of the auxiliary by unifying it with the variable θ-role; in this case (where *beaten DP*

[75] Clearly the choices for the elements are not arbitrary, i.e., the lists are to some extent related to one another lexically as well by the syntax, but this plays only a minor role in the following discussion.

[76] N.B. internal argument first (i.e. on top), for the reasons suggested above.

[77] Where the underline indicates a variable. This formalization, although clearly inspired by Prolog unification, is only given for the sake of concreteness to make the following discussion more specific and is not intended *per se* to make any theoretical claim about unification as a cognitive process.

[78] Of course to be more accurate the Case-list is a function of the root lexeme and the inflectional morphology: for example, *beat-* itself is marked only for accusative, *-s* (or rather [+Tense]) adds nominative, and the participle deletes the accusative.

[79] A misnomer, as should be clear from the discussion herein.

[80] That is, after unification its Case-list will also include the incomplete specification $<+ - \ldots \_C \ldots>$, for example.

[81] Exactly why *have* has this property is not clear. Cf. Verngaud (1985) for a suggestion, based among other things, on clitic climbing in Romance languages. See also discussion below.

is embedded under *have*), this will be the external θ-role of *beaten*, the internal one already having been bound to the DP and popped from the list. If there is no object present, unification producing the constant in the Case-list of *beaten* makes it a Case assigner, and hence its internal θ-role has the usual status of a true syntactic variable. If embedded under a form of *be*, whose Case-list presumably is something like (QC, Nom.), where QC = 'quirky case' whose morphological realization varies with language, the assignment is incompatible to an overt NP governed by the participle[82]; if there is no object, the "next" θ-role is unified upward, and in this case the internal one still heads the stack. Note that unification with the "next" θ-role is via the specifier position, since realising the internal argument in this position blocks the unification with the θ-position of the upper verb, as pointed out in (43) above.

Let us now look at the small-clause/ECM structures with *believe* more carefully. Suppose we have a projection which assigns no Case to its Specifier, like a *to*-projection [*to have left*] or a participle *beaten*, with a DP to their left under *believed*. If we unify the DP *John* with the projection of *to* forming [*John to have left*] or [*John beaten*] by the same mechanism, then the specification of the *to*-projection or $VP_{[+part]}$ will also contain a variable on its Case-list and be a Case-requiring element in the sense above. Again, the participle *believed* is not a Case-assigner *per se*, and combining/unifying the *to*-P or VP with *believed* adds the variable _C to *its* Case-list. If the projection of *believed* is governed by *have*, Case assignment to the DP in the Specifier of the *to*-$P/VP_{[+part]}$ follows by unification, giving us

(46) a. They had believed John to have left.

b. They had believed John beaten.

(with John taking the *internal* argument of *beaten* in the latter case) as desired. Note that we do not get the ungrammatical

(47) * They had believed [beaten John]

as only *had* has the property of Case-marking the internal argument of its complement, as noted above.[83] Hence *this* Case assignment cannot be applied recursively. Note again that if the Specifier position is blocked, as in (46b), then as in (43) the external θ-role is inaccessible, indicating that the unification for θ-linking is with the Specifier position. Implicit in this discussion, as in the case of Spec(COMP) discussed in §3.3.1 and $Spec(V_{[+part]})$ in (43) above, is that the specifier position functions as an escape-hatch, and seems to have structural status independently of a particular Case or θ-role assignment. That is, in the normal auxiliary cases as well as here, we need to assume that the Specifier position plays a crucial role in the linking unifying of the θ-positions.

There is one more case we need to consider, namely

(48) a. * John had been beaten the dog.

b. * John had been the dog beaten.

c. * John had the dog been beaten.

[82] Precisely why this cannot be assigned to a DP under a participle is a serious, and unsolved, problem, as noted in footnote 69. For example, QC in Russian may be realized as instrumental, *On byl studentom* = 'he was a student', but **On byl uviden studentom* = 'He was seen a student' is out in the intended sense, just as in English, etc. Cf. further discussion of "QC" cases below. Suggesting that BE "weakly" assigns this case just displaces the problem.

[83] Cf. example (45) above. While some speakers find (46b) more felicitous with *believe* than others, the examples with the internal argument realised *in situ* are completely unacceptible.

i.e., why the case assignment mechanism cannot be iterrated via the *been* to assign Case to *the dog*, in (a) and (b) or directly in (c). We can probably rule out the first case due to the exceptional assignment properties of *have* discussed above as suggested for (47). For (b) and (c), the only θ-role available for *John* is the "benefactive" one, as noted above. R. Huybregts (*pers.comm.*) suggests that what may be going on in (48c) is that in the benefactive θ-role (as well as the "causative" θ-role: *John had the garbage taken out*) the complement must be a "true" adjectival complement: cf.

(49) a. He still had the theorem unproved by 5:00.

b. ? They had all the bunkers indestructible before the air-raid that afternoon.

where *unproved* can only be an adjective. Hence, while the participle of ordinary verbs "counts" as a true adjective, that of *be* doesn't:

(50) a. A man unknown to the police ...

b. A man beaten by the police ...

c. * A man been seen by everyone ...

How this works out in terms of the feature specification is not entirely clear, but it seems promising. This leaves us with the case of (48b) where we suggest that the copula always requires a complement which can be predicated of its Spec-position, either as an identity or linked to the unsaturated argument position of a predicate; since the external θ-role in (b) is now unavailable, the latter is impossible and the former produces nonsense, hence the Spec(*be*) gets no interpretation at all and is out by some variant of the Principle of Full Interpretation (cf. Chomsky 1986a).

If we can overcome the problems noted with the peculiarities of the QC of *be* and the ability of *have* to address the internal Case position, we then have a unified account of the A-chain phenomena. This reinforces the conclusion drawn earlier: the external argument of the past-participle, if not available for θ-raising by having the internal argument discharged by exceptional Case-marking from above, is truly unavailable, as suggested in the previous section.

Noting this, let us return to a problematical case, namely that of *believe* in English. In the two preceding sections, we have claimed that all of the A-chain cases can be reduced to a mechanism which one needs in any case for assigning θ-roles to subjects of auxiliary verbs.[84] But then structure (51) is problematic:

(51) [John$_i$ was [e_i'' believed [e_i' to have left]]]

(extra e_i's omitted in the subordinate clause). The difficulty is that, in the case of passive, we have just suggested that the argument "available" for transmission is the "internal" one of the passivized verb, i.e., the "first" undischarged θ-role. In fact, believe has such a θ-role, and *it* passivizes normally in

(52) a. [The story]$_i$ was [believed θ_i by everone].

b. [That Jack left]$_i$ was [believed θ_i by no one].

c. It was [believed θ_i [that Jack left]$_i$].

Presumably we can maintain the usual line on examples like these, where the passive in (52a) is obligatory because of non-assignment of Case, and in (52b) it is optional (cf. 52c)

[84] Cf. discussion of a similar suggestion in Williams (1987); his proposal is subject to the same problem discussed in this subsection, as far as I can see.

since there is no Case requirement for C-projections (if anything, Case-resistance). However in (51), the internal θ-role is assigned to the complement; it cannot passivize itself with a phonetically present NP subject:

(53) * [John to have left] was believed

presumably for the same reasons (non-assignment of Case to its specifier) that small clauses and the dative cannot.[85] The difficulty is to suggest the mechanism by which the argument (θ-role) of the specifier of the *to*-projection takes precedence as the "available argument" of *believed* over *its* external argument, which is presumably "available", the internal argument having been "popped" by being assigned to the complement XP. Note that it is not the same as the other cases, as they are *inherently* copula-like; what we would need to insure is that passing the complement's available θ-role is *in fact* the default option. We can see that this indeed seems to be the default from similar constructions used as $\overline{\overline{V}}_{[+part]}$ adjuncts as in

(54) $[_{DetP_i}$ the$_i$ man [e_i believed e_i to have left]]

where movement is implausible.[86] Such a stipulation is highly unsatisfactory, and we leave the discovery of a better solution to further research.[87]

3.2.3 Raising adjectives

There is one additional wrinkle: the so-called raising verbs, like *seem*, have only an internal θ-role, which is assigned to their complement; this can be seen by the fact that the complement clause cannot appear in subject position (cf. Olsen 1981), and from the fact that raising is possible, i.e. the NP moves from a θ-position to a Case-marked but θ-less position:

(55) a. It seems [that John will leave].

b. * [That John will leave] seems.

c. John$_i$ seems [e_i to have left]

However, in the case of adjectives, there does seem to be a θ-role assigned to the subject position:

(56) [That John will leave] is likely.

In fact, even a normal NP can occupy the subject position:

(57) [John's success]$_{\theta_k}$ is certain.

If this is an external θ-role, as it superficially seems, then two quesions arise: firstly, how is "raising" possible (under the traditional analysis) or θ-passing, as discussed in the previous sections, since the subject would seem to be a θ-position? Presumably the answer to this is

---

[85] Assuming that the structure of *John gave Mary a book* is

John gave [$_{PP}$ Mary [[$_P$ e] a book]]

as required by binary branching. Cf. Holmberg (1986) as well as Barss & Lasnik (1986) for empirical evidence for this structure. A potential problem is that some German speakers accept preposed ACI structures like *[Den Fritz ein Lied singen] hast du doch nie gehört!*

[86] Cf. discussion in Williams (1987).

[87] One might countenance a solution along traditional lines: suppose the external argument *is* assigned to *John*. But then the structure of (51) would be something like

* [John$_{ext}$ was [θ_{ext} believed [PRO$_{arb}$ to have left]]]

and this structure is ruled out by the PRO in a governed position. This is perhaps not so bad, as the deviance of the goverened PRO can be checked locally, in the sense of (34).

that the θ-role is no longer assigned to the subject position, because it has been absorbed by the complement clause. But then we need to ask how a superficially external θ-role, the same one as in *His success is certain*, can be assigned to a Complement position, as the clause clearly is internal to the VP (or rather AP):

(58) Certain to win, he isn't!

Suppose we assume that θ_k in (57) is in fact the internal θ-role which is assigned to the complement, and that it *is* the same one as in

(59) Fred$_{\theta_j}$ is certain [that John will win]$_{\theta_k}$.

The cases where this θ-role appears in the subject position, as in (57), are then just like the passive and auxiliary verb cases (especially *believe*) discussed earlier: if there is no complement, then the "first" available θ-role is the internal one as before, and *it* is the one which will appear in the subject position; the external one, i.e., θ_j in (59), is unavailable, and in this case cannot be recovered by a by-phrase, since it is evidently not agentive. In (59), however, the internal one has been discharged, so the external one is available to be passed up, just as in cases where the internal θ-role is taken by an NP Case-marked with the dummy preposition *of*:

(60) John is proud of his mother.

The advantage of this analysis is that we can now see why the raising *verbs* cannot take sentential subjects: there is no way for the θ-role to be passed on to the subject position, whereas in the adjectival raising cases, the auxiliary can do so. That is, *seem* assigns *its* internal θ-role to the sentential complement, and its subject (Spec) can either be a dummy *it* or a subject with the θ-role of its complement, as in the cases of the auxiliary verbs discussed earlier, but there is no *syntactic* mechanism to allow its internal θ-role to be realised externally. (Cf. discussion below) In the adjectival cases, however, this is precisely what the copula does, namely passing up a θ-role of its complement.

Furthermore, we now have a simple explanation for the difference between two other well known classes of adjectives[88]:

(61) a. [That John will leave] is likely

b. [That John will leave] is probable

c. John$_i$ is likely [t_i to leave].

d. * John$_i$ is probable [t_i to leave].

If we assume that *likely* has an internal θ-role, like *seem*, and *probably* a truly external one, then the paradigm in (61) follows immediately as the subject position is already assigned a θ-role directly in the *probable* case, but not in the case of likely.

An apparent problem is that it seems impossible to prevent

(62) * [His mother]$_i$ is proud (e_i)

meaning the same thing as (60), i.e., being its "passive" version. Whatever prevents passing this θ-role is evidently idiosyncratic, as *certain* also allows an *of*-complement:

(63) John is certain of Fred's success.

Let us look at the paradigms more carefully: in the case of *verbs* we have, in addition to transitive/intransitive pairs like *eat*, transitive/ergative pairs like *grow*:

[88] Cf. discussion in Chomsky (1986b) where he refers to it as "a rather marginal lexical property"...

(64) a. θ_{ext} eat θ_{int} b. θ_{ext} grow θ_{int}
c. θ_{ext} eat d. grow θ_{int}

where in (d) there is presumably no Case assigned to the internal θ-role, and hence no external θ-role, causing the internal argument to be realized in the external position.[89] If we look at the full paradigm of *proud* and *certain* we see that these two cases are parallel, with *proud* like *eat* on the one hand:

(65) a. John is proud [that he won] / [of his mother].
b. John is proud.

and with *certain* like *grow* on the other:

(66) a. Jack$_{\theta_{ext}}$ is certain [that Fred left]/of [Fred's success]$_{\theta_{int}}$
b. It is certain [that Fred left]$_{\theta_{int}}$/ * of Fred's success.
c. [That Fred left]/[His success]$_{\theta_{int}}$ is certain.

That is,

(67) θ_{ext} proud θ_{int} θ_{ext} certain θ_{int}
θ_{ext} proud certain θ_{int}

What is different, and interesting, in the case of the Adjectives, is that we find some additional evidence as to *what* is specified in the lexicon: that Burzio's generalization does in fact seem to be at work. If we assume *of* to be a Case marker and allow the NP to realize the internal argument, we see that by (66b) that the [P NP] realization of the internal argument is impossible just when there is no external argument – in this case, however, the internal argument *can* be realized as a clause, which does not require case, and furthermore can also appear, optionally, in the Specifier position mitigated by the copula as in (66c).

Hence we can reduce the differences in the various classes to the same mechanism which we need to specify the ergative verb pairs in the lexicon, namely the presence or absence of Case/*of*, strengthening Burzio's generalization. In fact this makes it possible to express the intransitive/ergative distinction elegantly and succinctly: in the former case the internal *θ-role* is not specified, in the latter case, the internal *Case*.[90]

3.3 The status of Adjunct phrases

In this section we would like to ask whether adjunct-phrases undergo "movement" (or, *mutatis mutandis* are linked to an empty category) in the sense usually understood in Generative Grammar. By adjunct-phrase we mean non-arguments (modifying expressions), not argument phrases which arguably have been Chomsky-adjoined to another maximal category.[91]

[89] We will not go into the controvery about whether the actual "raising" takes place in the syntax or lexicon; cf. Napoli (1988) and references therein: it would seem that the elements are *marked* in lexicon but θ-passed in syntax.

[90] Returning briefly to the case of *seem*, we note that since the removal of Case (or the equivalent feature for sentential complements) in the ergative construction is a lexical process, it need not apply across the board. Why the past-participle of *seem* (*It had seemed true at the time*) cannot be used in a *syntactic* passive (**That Fred left was seemed*) is apparently related to its status as a modal (under the analysis here) – note that in languages where other modals, unlike English *do* have participle forms, they nevertheless cannot passivize: **Die schwierigsten Sonanten van Beethoven spielen wurde von/durch Fritz gekonnt*. This underscores the crucial role of the auxiliaries in θ-role passing.

[91] Such as *meinen Sohn* to *S* in 'Gestern hat [$_S$ [meinen Sohn]$_i$ [$_S$ ein Hund e$_i$ gebissen]]'; cf. example (27) on page 17. Cf. discussion in Saito (1985) and Thiersch (1985), among others.

How are adjuncts (in this sense) licensed? A minimum assumption is that modifier adjuncts are basically unsaturated predicates which have to be licenced by having their missing argument linked to an appropriate element. Recall from the previous section the analysis of auxiliary verbs, in which the unsaturated θ-role of the preposition or adjective[92] passed on the the specifier position of the *auxiliary* (or copula); cf. example (31), repeated here:

(31) $[_{\overline{\overline{V}}_{[+fin]}}$ John$_i$ is [$_{PP}$ θ_i [in the kitchen]]]

If the constituent is not governed by such a head (e.g., copula) and the specifier is not discharged by being assigned Case, as in a small clause construction, then it is available for linking (in the default case) as a modifier of the head. Hence in the PP [θ_i *in the kitchen*], the argument position θ_i which was linked to the subject in *John is in the kitchen* will be linked to the referent of the NP (i.e., DetP) in *the man in the kitchen*:

(68) [$_{DP_i}$ The$_i$ man [θ_i in the kitchen]]

and similarly in *a patient sick with the flu*, *a boy proud of his mother*, *the protesters beaten by the police*. Note that the θ-position of the adjunct need not be phonetically empty, as in the case of relative clauses, where it may be filled by an overt WH-pronoun. They may also under certain circumstances be linked to either the complement or specifier (cf. *Jack ate the meat quickly/raw/naked*).[93]

The exact details of this linking process are still unclear; cf. for example Travis (1987) for a discussion of the placement of adverbial expressions. For our purposes, it is sufficient to note that some sort of linking like this is the descriptive minimum for interpretation and hence, by the "Principle of Full Interpretation" (Chomsky 1986a), for the licensing of *in situ* adjuncts.

3.3.1 Adjuncts in the *Vorfeld*

While one might want to suggest that the scope, or domain, for this interpretive linking is the "host" projection, and that cases of $\overline{\overline{V}}$ (i.e., $\overline{\overline{INFL}}$ or S) interpretation are generated by extraposition similarly to complements, as in the complement of the subject in *The trial began yesterday in Southern Yemen of the five terrorists who* ..., there seems to be reason to believe that the adjunct cases are not, in fact, "movement" but rather interpretation within a local domain, e.g. $\overline{\overline{COMP}}$. Let us look more closely at the question as to whether Adjuncts in the Vorfeld, i.e. Spec(COMP), such as *gestern* in

(69) Gestern ist er in die Oper gegangen.
yesterday is he in the opera gone
'He went to the opera yesterday.'

are "moved" there, i.e., receive their interpretation by being bound to an independently licensed empty category, or whether they are interpreted *in situ* in this position as well.[94] If Adjuncts in the Vorfeld *do* bind a position in the Mittelfeld, then their presence in the

[92] Its "specifier position"; cf. more detailed discussion in §3.2 (A-chains).

[93] We would like to suggest that (in the *default* case) they are linked to the referent of the head of their projection, as there seems to exist a clear preference for linking directly to the "host-projection" if this is an option at all.

[94] This issue was brought up, in part, by work on the parser, as *arguments*, being independently licensed by a head, can simply be unified, whereas *adjuncts* involve building-up of an extra layer of phrase-structure, and displaced adjuncts involve *retrospective* structure-building in a partially completed structure. (This is independent of the so-called "attachment problem", much discussed in the parsing literature, and about

Vorfeld is presumably licensed in a similar manner to Arguments, by being properly bound to a licenced base position, which has a certain conceptual elegance to recommend it[95].

However, in addition to the parsing consideration of retrospective structure building noted in footnote 94, there is also a certain redundancy; if we already have the interpretive mechanism *and* the domain of interpretation can be $\overline{\overline{C}}$, the adjunct will be independently licenced, and it is not obvious that it needs to be bound to a empty category. What are the consequences if we assume that, in the normal case, adjuncts in the Vorfeld are licensed by interpretation? First of all, it has often been noted in the linguistics literature that adjuncts behave rather differently with respect, for example, to the Empty Category Principle. That is, they are subject to more stringent locality conditions[96], as in the canonical example

(70) a. ? What did John wonder how Fred fixed?

b. * How did John wonder what Fred fixed?

where *how* is interpreted in the subordinate clause.

One might well ask, *why* it should be the difference in government that causes the contrast; since government is presumably necessary to license an *argument* position on the basis of (obligatory) Case and θ-assignment, and adjuncts are rather licensed by interpretation, we might suspect that the difference results from the difference in licensing. That is, *arguments* in the Vorfeld are licensed by being bound in an A-bar chain to an independently licensed position, whereas *adjuncts* are interpreted in the domain $\overline{\overline{C}}$.

We can gain some insight into (70) by looking at the Spec(COMP) in German; first, containing an independently licensed element is not the only way in which Spec(COMP) (the Vorfeld) is licensed *as a position* — there seems to be some justification for the Vorfeld itself to be independently licensed,[97] at least in main clauses: First of all, we note that it is not a Case position itself, and the *es* in the following arguably has no θ-role, since it is obligatory in the main clause Vorfeld, but ungrammatical in the subordinate clause, where, being in the Mittelfeld, it would presumably be in a true argument position:

(71) a. Es wurde getanzt. (*cf.* *..., daß es getanzt wurde)
it was danced
'There was dancing.'

b. Es dünkt mich, ... (*cf.* *..., daß es mich dünkt, ...)
it thinks me
'I think that ...' (*cf. archaic* 'Me thinks')

---

which we have nothing more to say. Cf. Marcus, et al. (1983) for a promising suggestion to reduce the structure-building problem to the interpretation/attachment one.)

There might be a possible argument for retrospective structure building (interpretation via an empty category) if we were to capture scope differeces of the sort

i. The *tall* student with red hair ...

ii. The tall student with *red* hair ...

as differences in bracketing: [AP [N PP]] vs. [[AP N] PP]. The same readings seem to to be available in *We saw the tall student yesterday with read hair.* The question is whether the difference between (a) and (b) ought to be captured by differences in phrase structure at (S-structure); there are similar differences in focus where no difference in constituent structure is plausible at S-structure, e.g. in the German Mittelfeld. Cf. Höhle (1982).

[95] Thanks to Teun Hoekstra for getting the ball rolling by challenging this during a lecture.

[96] Cf. discussion in Lasnik and Saito (1984) and references therein, and subsequent discussion in Chomsky (1986b), as well as the following discussion.

[97] Indeed, we ought to suspect this in a number of cases; otherwise one could create enormous structures by vacuously Chomsky-adjoining. Discussion of a suggested solution would take us beyond the scope of this article.

c. Es kamen gestern zwei Männer. (*cf.* *Gestern kamen es zwei Männer.)
it came yesterday two men
'There came two men yesterday.'

mich dünkt,

Secondly, arguments in the Mittlefeld normally may not be omitted:

(72) * Gestern habe ich $e_{[acc.obj.]}$ nicht gesehen.
yesterday have I *e.c.* not seen
*'*I didn't see – yesterday'*

The one exception is when the Vorfeld is empty:

(73) Habe ich überhaupt nicht gesehen!
have I at all not seen
'I didn't see (it) at all!'

Sentences like (73) presumably have the analysis

(74) [OP_i habe ich e_i überhaupt nicht gesehen][98]

where OP_i is an empty operator binding the trace in argument position. That is, in order to count as a declarative clause, the Vorfeld, Spec(COMP), *must* be filled, even if only with a dummy phonetic item as in (71) or with an empty operator, as in (73). Hence we need to assume some sort of pseudo-argument status for the Specifier of COMP.[99]

If the specifier of COMP is a "potentially" available position, by virtue of its *structural* status, then we have an explanation for the grammatical example (75a), related to (70), repeated as (b-c) below:

(75) a. How$_i$ did Frank think [e_i that Sam fixed the car]?
b. ? What did John wonder how Fred fixed? [= 70a]
c. * How did John wonder what Fred fixed? [= 70b]

The adjunct may (exceptionally) be interpreted within the lower $\overline{\overline{C}}$ by virture of being bound to the *independently licensed* Specifier of COMP.[100]

The contrast in (70) follows immediately: as just noted above, the Specifier of the C-projection seems to be implicitly available on independent grounds, perhaps licensed by its

[98] See Huang (1984) and Ross(1982) for a more detailed discussion of this construction.

[99] Intuitively, we would like to distinguish between a true argument position which is assigned a θ-role directly, and the Spec-position of some constituent $\overline{\overline{X}}$, which is interpreted as fulfilling *some* θ-role, either directly (from the $\overline{X}$) or indirectly (from the complement of the X^0). In this respect Spec(COMP) is behaves similarly to Spec(INFL), i.e. Spec([+V,+Tense]):

i. Fred kicked the bucket.

ii. Tomorrow is a good time to leave.

iii. Sam seems tired.

[100] Unless, of course, one would try to develop an analysis by which the Adjunct *how* would link to the index of the lower $\overline{\overline{C}}$, which would then somehow be related to the V[+tense] ~ INFL projection, but this seems unlikely in view of examples like

i. * How is it time [e_k for John to fix the car]?

in which the Spec of the adjunct phrase, e_k is already occupied by being linked as a modifier (cf. discussion below), and *how* cannot modify the adjunct internal V-projection. We will not pursue this here.

status as a "pseudo-argument" position of the Head.[101] Hence in (75a) the adjunct (i.e., the empty category) may be interpreted in the second clause. If we view A-bar binding of arguments, however, as a search (either through a list or sections of a tree), then there is a possibility of recovering the displaced *argument* — i.e., unifying it with an already existing (i.e., licensed) empty category in (75b), even when Spec(COMP) is occupied; but the displaced *adjunct* in (75c) is impossible to interpret — Spec(COMP) is filled, and otherwise its domain of interpretation is its own $\overline{\overline{C}}$.[102] Hence we do not lose an explanation for the standard "ECP" case; it follows from already needed licensing conditions.

Note that this will also provide some basis for the contrast in judgements between subject and adjunct violations: while there is an assymetry between subject and object violation (the standard ECP effect contrast):

(76) a. ? What did Fred wonder who fixed ?

b. *? Who did Fred wonder what fixed?

the corresponding adjunct extraction violation (75c) is not just unacceptable, like (76a,b), but is *uninterpretable*; the reading is simply not available.

Note also that assuming that the specifier position of adjuncts is used to link for modification will also account for the island properties of adjunct phrases, the specifier "escape-hatch" is not available — arguments can be extracted with difficulty, with the usual subjacency violations, but *adjuncts* cannot be extracted at all. Recall footnote 100.

3.3.2 The "*by*-phrase" revisited.

We are now in a position to suggest a solution for the dangling problem of the *by*-phrase in passive, noted in § 3.2.1, namely, if we assume that the external θ-role is unavailable outside of the local domain (bottom-most V-projection), (i.e., taking the stack-model proposed in § 3.2.2 literally), we should expect the external θ-role to be *truly* unavailable in passives and in particular unable to be realized by the by-phrase.

There is considerable discussion of this problem in the literature,[103] but we would like to ask, does the account sketched above suggest a particular solution as being the correct one? What could it mean in terms of this acccount for the agent θ-role to be "picked-up"? If transitive prepositions are, as suggested, two-place relations, then *the boy was beaten by the police* has the following structure

(77) [the boy]$_{int}$ was [beaten: $((\theta_{ext})\,\theta_{int})$] [PP θ_k by the police]

To what is the the external argument θ_k of the PP linked? According to the interpretation just sketched for adjuncts, it ought to be linked to the referent of the participle *beaten*. If it is the preposition *by* which assigns its complement the agent θ-role, which just coincidentally for some verbs happens also to be the external θ-role, cf.

(78) a. the window broken by John

b. * the arm broken by John [* *in the intended sense*]

then we can claim that in fact the external θ-role is *always* completely unavailable in these cases; i.e., it is by virtue of θ_k being linked to the referent of *beaten* and hence modifying it that the two cases happen to have the same meaning, just as in *The burgler stabbed her with*

[101] As it was in our proto-type parser.

[102] Actually the matter is a bit more complicated than this in the case of coordiate structures; cf. discussion in Thiersch (1988).

[103] Cf. Höhle (1978), esp. chap.7 and further discussion in Zubizaretta (1986).

a knife the burgler presumably has the knife without assuming that we have a *syntactic* link between the two as in [$_{\text{DetP}_i}$ *the$_i$ burgler* [θ_i *with a knife*]].

Thus, if the assignment really does work like a stack, as suggested in § 3.2.2, we arrive at a uniform system for interpreting both specifiers and adjuncts relying on only the most minimal assumptions[104], and we are able to retain a plausible analysis for the canonical difference in the behavior of adjunct versus arguments with respect to extraction, linking it to the difference in licensing.[105] This means that we can restrict our attention to cases of A-bar movement.

4 Conclusion: Constraining move-α

Let us take stock: Under a derivational interpretation, the inherent complexity of a principles and parameters based theory employing distinct levels of represention almost inevitably leads to computational intractability. Under a non-derivational interpretation, however, the structure corresponding to all the levels can be constructed simultaneously in an incremental fashion. The success of principles-based parsing then largely depends on the procedural as well a theoretical locality of the application of the modules. We have been pursuing two lines of argumentation: first, that one of the problems of "traditional" rule based parsers is solved by the uniform constituent hypothesis – one is not building structure by matching *arbitrary* phrase structures, for example, given by (CF) rules like $A \rightarrow BCD\ldots$, but rather that one is in fact making decisions involving two elements, a head(-projection) and satellite on the basis of straightforward and local criteria. An approach to A-chain relationships has been suggested which not only does not rely on a separately generated D-structure for licensing,[106] but is even more strictly local than the interpretation usually required for A-chains, namely the structure (34) in § 3.2, in which (aside from adjuncts of course) the discharging of the θ-role in the higher projection is mitigated by the head V_k, i.e., is *required* by the upper head. Hence in the structure noted earlier (§2, (24)) repeated here,

(24)	a. Who do you believe [*e* to be funny]	*referential expression*
	b. Jack is believed [*e* to be funny]	*anaphor*
	c. Jack tries [*e* to be funny]	*PRO ("pronominal anaphor")*

we know the nature of the empty category, in the best case, immediately upon seeing the governing verb *believe*, *believed*, or *try*, and in the worst case, as soon as we see the auxiliary *have* or *be*, and hence at that point we know whether or not we need to worry about binding the argument position.[107] Adjuncts (in the modifier sense), we have suggested, do not add to the proliferation of ambiguous *structure*, as they are never moved, so there is no question of placing and replacing empty categories.[108]

This only partially solves the serious problem, noted above in §2, of how to constrain the multitude of possible structures early on in the parse; the primary source of this was

[104] Eliminating such mysterious concepts as "indirect θ-marking" of adjuncts (Cf. Koopman 1984).

[105] Additional evidence for licensing adjuncts in the Vorfeld by interpretive linking within $\overline{\overline{\text{C}}}$, rather than syntactic binding to a position, can be deduced from some exotic German data, the Subjektlücke (SL) construction, described in Höhle (1983). It would take us too far afield to discuss this here; cf. Thiersch (1988).

[106] Which (cf. §2 (19) and footnote 28) is, strictly speaking, unnecessary once one adopts trace-theory of movement.

[107] In a left-right parse of course we have this information before we encounter the gap, but then the analagous problem arises in head-final languages.

[108] The so-called attachment problem of course remains, as noted above, but this, while a *parsing* problem is not a *syntactic* one; that is, the solution must to some extent rely on semantics and discourse interpretation.

the enormous number of possible combinations of various adjunctions through application of move-α. That is, even if we limit "true" movement (A-bar chains) to movement from properly licensed Specifier and Complement positions,[109] we are overwhelmed with multiple structures. The villain is, of course, scrambling: if we view scrambling (and heavy NP-shift, extraposition, and a multitude of other constructions) as binding an empty category,[110] then we are still back at Square One if we allow move-α to apply freely because of the multiplicity of possibilities. Unfortunately, sketching a proposal to restrict this would take us beyond the scope of this short article, and we leave this to forthcoming publications.

References

[1] Abney, S. (1985) "Functor Theory and Licensing Conditions: Toward elimination of the base," unpubl. ms., MIT, Cambridge, Mass.

[2] Aho, A. V. and J. D. Ullman (1972) *The Theory of Parsing, Translating and Compiling,* Prentice Hall, Eaglewood Cliffs, N.J.

[3] Barss, A. and H. Lasnik (1986) "A Note on Anaphora and the Double Object Construction," LI 17.2, p.347-54.

[4] Bennis, H. and T. Hoekstra (1985) "Parasitic Gaps in Dutch," NELS 15, Brown University, *Proceedings,* GLSA, U. Mass./Amherst.

[5] Burzio, L. (1981) *Intransitive Verbs and Italian Auziliaries,* Ph.D. Diss. MIT, Cambridge, Mass.

[6] Burzio, L. (1986) *Italian Syntax,* Reidel, Dordrecht.

[7] Cann, R. (1986) "Heads without Bars: a Theory of Phrase Structure," unpubl. ms, U. of Edinburgh.

[8] Chomsky, N. (1970) "Remarks on Nominalization," in R. Jacobs and P.S. Rosenbaum (eds.) *Readings in English Transformational Grammar,* Ginn & Co., Waltham, Mass.

[9] Chomsky, N. (1975) *Reflections on Language,* Pantheon, N.Y.

[10] Chomsky, N. (1981) *Lectures on Government and Binding,* Foris, Dordrecht.

[11] Chomsky, N. (1982) *Some Concepts and Consequences of the Theory of Government and Binding,* MIT Press, Cambridge, Mass.

[12] Chomsky, N. (1986a) *Knowledge of Language,* Praeger, N.Y.

[13] Chomsky, N. (1986b) *Barriers,* MIT Press, Cambridge, Mass.

[14] Chomsky, N. (1988) "Some Notes on Economy of Derivation and Representation," unpubl. ms., MIT, Cambridge, Mass.

[15] Cichocki, W. (1983) "Multiple WH-Questions in Polish: a Two Comp Analysis," *Toronto Working Papers in Linguistics,* 4, p.53-71.

[16] Comorovski, I. (1986) "Multiple WH-movement in Romanian," LI 17.1, p.171-77.

[109] I.e., not "positions" which are not realized by virture of failing to have been assigned some feature such as Case or verbal inflection, such as the internal argument of the verbal particple in passive.

[110] and it's difficult to know how else to view most of these constructions in spite of protestations to the contrary in the literature ...

[17] Csepluch, H. (1988) "Kasusmorphologie und Kasusrelationen: Überlegungen zur Kasustheorie am Beispiel des Deutschen," *Linguistische Berichte* 116, p.275-310.

[18] Evers, Arnold (1975) *The Transformational Cycle in Dutch and German,* Ph.D. Diss., U. Utrecht.

[19] Felix, S. (1983) "Parasitic Gaps in German," U. Passau, *Working Papers.*

[20] Fukui, N. and M. Spaes (1986) "Specifiers and Projections," unpubl. ms., MIT, Cambridge, Mass.

[21] Gazdar, G., G. Pullum, and I. Sag (1982) "Auxiliaries and Related Phenomena in a Restrictive Theory of Grammar," *Language* 58.3, p.591-638.

[22] Haegeman, Liliane and H. van Riemsdijk (1986) "Verb Projection Raising, Scope, and the Typology of Rules Affecting Verbs," LI 17.3, p.417-66.

[23] Höhle, T. N. (1978) *Lexicalische Syntax,* Niemeyer, Tübingen.

[24] Höhle, T.N.(1982) "Explikation für 'nomale Betonung'," in Abraham, W. (1982) *Satzglieder im Deutschen,* Narr, Tübingen.

[25] Höhle, T. N. (1983) "Subjecktlücken in Koordinationen," unpubl. ms., U. Köln.

[26] Holmberg, A. (1986) *Word Order and Syntactic Features,* Ph.D. Diss. U. Stockholm.

[27] Huang, C.-T. J. (1984) "On the Distribution and Reference of Empty Pronouns," LI, 15.4, p.531-74.

[28] Huybregts, R. and H. van Riemsdijk (1985) "Parasitic Gaps and ATB," NELS 15, Brown University, *Proceedings,* GSLA, U. Mass./Amherst.

[29] Johnson, M. (1987) "The Use of Knowledge of Language," unpubl. ms., MIT, Cambridge, Mass.

[30] Karttunen, Lauri (1986) "D-PATR. A Development Environment for Unification-Based Grammars," SRI/CSLI, Stanford, Calif.

[31] Kayne, R. (1984) *Connectedness and Binary Branching,* Foris, Dordrecht.

[32] Koster, J. (1978) *Locality Principles in Syntax,* Foris, Dordrecht.

[33] Koster, J. (1987) *Domains and Dynasties,* Foris, Dordrecht.

[34] Kolb, H.-P. (1989) "Op principes gebaseerde ontleding," in A. Neijt & D. Bakker (eds.) *Computerlinguistiek: een overzicht in artikelen,* Foris, Dordrecht.

[35] Koopman, H. and D. Sportiche (1988) "Subjects," unpubl. ms. UCLA, Los Angeles, Calif.

[36] Lasnik, H. and M. Saito (1984) "On the Nature of Proper Government," LI 15.2, p.235-89.

[37] Lenerz, J. (1977) *Zur Abfolge nominaler Satzglieder im Deutschen,* Narr, Tübingen.

[38] Marcus, M. (1978) *A Theory of Syntactic Recognition for Natural Language,* Ph.D. Diss. MIT, Cambridge, Mass.

[39] Marcus, M., D. Hindle, and M. Fleck (1983) "Talking about Talking about Trees," *Proceedings of the 21st Meeting of the Association for Computational Linguistics.*

[40] Moortgat, Michael (1987) "Parsing Lambek Categorial Grammars," INL Working Papers 87-05.

[41] Muysken, P. (1982) "Parameterizing the Notion 'Head'," JLR, no.2.

[42] Napoli, D. (1983) "Review of Burzio: *Italian Syntax*," *Language* 64.1, p.130-42.

[43] Olsen, S. (1981) *Problems of* seem/scheinen *Constructions and their Implications for the Theory of Predicate Sentential Complementation*, Niemeyer, Tübigen.

[44] Pesetsky, D. (1982) *Paths and Categories*, Ph.D. Diss. MIT, Cambridge, Mass.

[45] Pollock, J.-Y. (1989) "Verb Movement, Universal Grammar, and the Structure of IP," LI 20.3, p. 365-424.

[46] Riemsdijk, H. van (1987) "Movement and Regeneration," unpubl. ms., Tilburg U.

[47] Ristad, E. (1988) "Locality in Computation and Language," paper presented at the 11th GLOW Colloquim, Budapest.

[48] Ross., J. R. (1982) "Pronoun Deleting Processes in German," paper presented at the LSA Annual Meeting, San Diego, Calif.

[49] Saito, M. (1985) *Some Assymetries in Japanese and their Theoretical Implications*, Ph.D. Diss. MIT, Cambridge, Mass.

[50] Sportiche, D. (1988) "Conditions on Silent Categories," unpubl. ms., UCLA, Los Angeles, Calif.

[51] Steedman, M. (1987) "Combinators and Grammars," in. R. T. Oehrle, E. Bach and D. Wheeler *Categorial Grammars and Natural Language Structures*, Reidel, Dordrecht.

[52] Stowell, T. (1981) *Origins of Phrase Structure*, Ph.D. Diss. MIT, Cambridge, Mass.

[53] Steen, G. Van der (1987) *A Program Generator for Recognition, Parsing and Transduction with Syntactic Patterns*, Proefschrift (Doctoral Diss.) U. Utrecht.

[54] Thiersch, C. (1985) "Some Notes on Scrambling in the German Mittelfeld, VP and X-bar Theory," unpubl. ms., U. Conn & U. Köln

[55] Thiersch, C. (1988) "A note on the Status of Adjuncts," unpubl. ms., U. Tilburg.

[56] Thiersch, C. (1989; *forthcoming*) "Het (hulp-)werkwoord Systeem," in *Syllabus: Grammatische Analyse*, Chap. 11, U. Tilburg; English version in preparation, probably to appear in proceedings of 1990 Geneva Workshop on G/B Parsing.

[57] Toman, J. (1982) "Aspects of Multiple Wh-Movement in Polish and Czech," in R. May and J. Koster, eds. *Levels of Syntactic Representation*, Foris, Dordrecht.

[58] Tomita, M. (1986) *Efficient Parsing For Natural Language. A Fast Algorithm for Practical Systems*, Kluwer, Dordrecht.

[59] Travis, L. (1984) *Parameters and Effects of Word Order Variation*, Ph.D. Diss. MIT, Cambridge, Mass.

[60] Travis, L. (1988) "The Syntax of Adverbs," in *McGill Working Papers in Linguistics: Proceedings of the IVth Workshop on Comparative Germanic Syntax*, McGill University, Montréal.

[61] Verngaud, J.-R. (1985) *Dépendances et niveaux de représentation en syntaxe*, Benjamins, Amsterdam.

[62] Williams, E. (1981) "Argument Structure and Morphology," LR 1.1, p.81-114.

[63] Williams, E. (1986) "A Reassignment of the Functions of LF," LI 17.2, p.265-99.

[64] Williams, E. (1987) "NP Trace in Theta Theory," *Linguistics and Philosophy* 10.4, p. 433-47.

[65] Zubizarreta, M.-L. (1986) *Levels of Representation in the Lexicon and in the Syntax*, Foris, Dordrecht.

Abbreviations:

JLR = *Journal of Linguistic Research*

LI = *Linguistic Inquiry*

LR = *The Linguistic Review*

ITK: P.O. BOX 90153 5000 LE TILBURG THE NETHERLANDS